The IEP Data Collection Intentions Scale (IDCIS): Scale Development and Validation for Intended Score Interpretation and Use in Early Childhood

A Dissertation
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Brenna Noland Rudolph

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. LeAnne Johnson, Advisor

August 2019

Acknowledgements

I am beyond grateful to a number of individuals who have supported me in various capacities throughout this process. As such, many acknowledgements are in order.

To my advisor, Dr. LeAnne Johnson who has allowed my interests to guide every aspect of this experience, while providing just enough support to keep me sweating! Your fearless leadership and high expectations have taken my learning to new heights.

To the rest of my committee members, Dr. Scott McConnell, Dr. Michael Rodriguez, and Dr. John LaVelle who have listened thoughtfully, provided fresh perspectives, and answered endless questions. Your guidance—both in and out of the classroom—has made this process more enjoyable than one might think!

To Kyle Nickodem, who played an instrumental role in my ability to pull this off. Not only are you crazy talented when it comes to analyzing data, but you have a knack for facilitating others' capacity and confidence related to statistics.

To the hundreds of ECSE teachers across the state who took time out of their busy schedules to participate in this study. While our relationship was brief (and maybe a bit one-sided!), this project would not have been possible without you. I hope to someday pay you back with new and improved practices that make IEP data collection more manageable.

To my fellow doctoral students, who over the past five years have become some of my most cherished friends. This has been a wild ride; one that I could not imagine going on (and surviving!) without all of you!

To my Mom, Renée for all you have done for me and my family to make this goal a reality. I love you!

Last, but most definitely not least, to my little Willa, the absolute best distraction on even the toughest of days. Your "No phone! Pocket!" orders always came at the perfect time, reminding me to take a break and fully embrace the greatest journey of all—being your mama! While completing this dissertation is close to the top, you are by far my proudest accomplishment.

Dedications

Without hesitation, I dedicate this dissertation to my husband, Arik who in so many ways has made these last five years possible. You truly are the best; and for that, I thank you!

Abstract

There is evidence to suggest a research-to-practice gap exists in regard to Early Childhood Special Education (ECSE) teachers' collection of data highlighting students' progress toward meeting their Individualized Education Program (IEP) goals and objectives (i.e., IEP data collection). Due to the negligible amount of research in this area in addition to the limitations present in the literature, however, it is unclear what factors are responsible for causing and maintaining this gap. Given that teachers are ultimately responsible for deciding whether and how to engage in IEP data collection, a focus on better understanding teachers' intentions to collect IEP data is a logical first step. With an emphasis on enhancing the measurement techniques employed in previous studies, this application of a cross-sectional survey design aimed to validate the intended interpretations and uses of scores resulting from administration of a newly developed scale—the IEP Data Collection Intentions Scale (IDCIS). Following survey completion by 368 ECSE teachers across the state of Minnesota, confirmatory factor analysis, item analysis, and item response modeling were performed to support scale development. Results indicated that following minor adjustments, the IDCIS can be used to produce precise measures of teachers' attitudes, subjective norms, self-efficacy, controllability, and intentions related to the collection of IEP data. Furthermore, the scores produced by IDCIS administration can be used to make valid and reliable inferences about teachers' levels of each construct in order to inform the creation and modification of future implementation supports, thus decreasing the gap between what is known and what is practiced in today's classrooms related to data collection.

Table of Contents

List of Tables

List of Figures

Chapter One – Introduction

An abundance of empirical evidence exists supporting the use of a diverse collection of instructional practices in special education. Additionally, the availability of various training modalities (e.g., educational journals, online webinars and modules, conferences, coaching) designed to increase educators' knowledge of these practices continues to grow. A large gap, however, continues to exist between what we know and what educators actually do (Cook & Odom, 2013; Gersten & Smith-Jones, 2001; Odom, 2009). Though it tends to receive less attention in the literature than other instructional practices, a gap between research and practice has been identified with respect to the practice of ongoing collection and use of data for the purposes of child-specific instructional planning (Brawley & Stormont, 2013; Cook et al., 1991; Fuchs, Fuchs, & Warren, 1982; Sandall, Schwartz, & LaCroix, 2004; Wesson, King, & Deno; 1984)—also referred to as progress monitoring.

Progress monitoring is not a new concept; the iterative process of frequent data collection and use to aid in the planning of individualized instruction has been described in the literature for several decades, dating back to the conception of Data-Based Program Modification (DBPM; Deno & Mirkin, 1977). Originating from Stan Deno's work at the University of Minnesota and created in response to the Education for All Handicapped Children Act (EAHCA; Public Law 94-142, 1975), DBPM has been defined as a "systematic method of individualizing educational plans for children with any kind of learning or behavioral problem" (Deno & Mirkin, 1977, p. 4). Deno and colleagues recognized that while many models of individualization had been proposed to facilitate teachers' ability to provide individualized instruction, no one system worked for all students across all contexts. As opposed to a specific tool or structured method of collecting and using data, DBPM included a set of fluid procedures that facilitated the evaluation of alternative solutions to *any type of problem* that *any type of student* with special needs might face in the classroom, not only resulting in more successful program modification, but also making its use more universal than previously proposed systems (Deno, 2016). In addition to DBPM, several other terms have been used by researchers and practitioners alike to represent the cycle of progress monitoring, including data-based decision making (e.g., Crum, 2009), data-driven decision making (e.g., Marsh, Pane, & Hamilton, 2006), data-based instruction (e.g., Jones & Krouse, 1986), and data-based individualization (e.g., National Center on Intensive Interventions, 2013).

Data Collection

While not emphasized in its name, the significance of *data collection* within the progress monitoring cycle is evident in the fundamental assumptions grounding DBPM: 1) individualized instructional programs are merely hypotheses in need of empirical testing; 2) time series designs, during which data are collected both systematically and frequently over time, are most appropriate for testing these hypotheses; 3) individual student performance must be tested; 4) educational "vital signs" must be identified in order to facilitate data collection; and 5) effective collection and use of data requires trained professionals (Deno & Mirkin, 1977). These assumptions are an important reminder that while ample evidence supporting the efficacy of various practices with students with disabilities exists, one cannot be certain of the success of any one practice in particular given the number of context-specific variables known to impact effectiveness. In order to ensure program effectiveness and student progress, therefore, educators must be trained in the ongoing collection of data highlighting students' performance of essential skills included in their Individualized Education Programs (IEPs). Though the data are meaningless unless used, data may only be used if collected. Therefore, the collection of IEP data—herein referred to as IEP data collection—is the focus of this study, given its sequential position in the process of progress monitoring.

Statement of Problem

Knowing that the systematic and ongoing collection and use of data has been linked to higher overall achievement in students with disabilities (Fuchs, Deno, & Mirkin, 1984; Mirkin, Deno, Tindal, & Kuehnle, 1982; Stecker & Fuchs, 2000) as well as more effective instructional planning (Fuchs, Fuchs, & Stecker, 1989; Stecker & Fuchs, 2000), the discrepancy between what is known and what is practiced in today's classrooms is concerning. Furthermore, adding to this concern is the fact that special educators are mandated to develop IEPs that include measurable goals in addition to a description of how student progress will be measured (IDEA, 2004). Development of the IEP, however, is not enough; progress toward goals must be measured (Christle & Yell, 2010).

While attention toward increasing the progress monitoring practices of special educators across the primary and secondary grades (K-12) is mounting, progress monitoring practices in early childhood special education (ECSE)—a field representing the primary service for preschool-aged children with disabilities and their families (Odom & Wolery, 2003)—has been comparatively overlooked. Though little is known about the collection of IEP data in ECSE, there is some evidence to suggest that ECSE professionals share the perspective that data collection is important for a variety of reasons including accountability, development of educational goals, and instructional planning; yet, a conflict exists between their self-reported beliefs about its importance and their actual practices (Brawley & Stormont, 2013; Ruble, McGrew, Wong, & Missall, 2018; Sandall et al., 2004). Given the importance of IEP data collection, a number of supports have been developed to facilitate this practice in ECSE, from structured frameworks for collecting and analyzing data to numerous data collection tools and various electronic technologies. Nevertheless, the sporadic, unstructured, and relatively subjective data collection methods continue to persist (Brawley & Stormont, 2013; Sandall et al., 2004).

Shedding some light on this inconsistency are ECSE professionals' reports of the many variables serving as barriers to their collection of data, most of which can be categorized as external factors impacting the feasibility of IEP data collection. The negligible amount of research in this area in addition to a number of limitations existent in the literature, however, have prevented a clear understanding of how ECSE teachers' reported beliefs, identified barriers, and IEP data collection practices are related. In addition to the lack of an organizing framework guiding and connecting individual research efforts, the inconsistent defining of constructs within and across studies and the questionable measurement techniques employed have made it difficult to accurately interpret the findings. Furthermore, because each individual teacher is ultimately responsible for deciding if and how to engage in IEP data collection, a lack of focus on better understanding teachers' motivation or commitment to collect IEP data has left many questions unanswered when attempting to reduce this research-to-practice gap.

Study Purpose and Research Questions

To assist in the creation of implementation supports designed to increase teachers' ability and commitment to persist with IEP data collection regardless of present barriers, it is imperative that we rely on a theory-driven approach to the identification, defining, and measurement of key constructs thought to impact one's data collection practices. The Theory of Planned Behavior (TPB; Ajzen, 1985), which is gaining more presence in educational research, has the potential to support our measurement and subsequent understanding of teachers' IEP data collection practices by focusing on a critical individual-level factor linked to implementation—teachers' intentions. As such, the purpose of this study was to evaluate the dimensionality and quality of the IEP Data Collection Intentions Scale (IDCIS), a scale grounded in the TPB and designed to collect data necessary to the creation of effective implementation supports tailored toward the specific contexts in which ECSE teachers work. In doing so, answers to the following three research questions were explored:

1. To what extent, if at all, does the IDCIS represent four distinct constructs, including teachers' attitudes, subjective norms, perceived behavioral control, and behavioral intentions toward IEP data collection?
2. Which items serve as quality indicators of each construct, such that valid and reliable inferences about teachers' IEP data collection intentions—regardless of level—can be made using the resulting data?
3. Does the Theory of Planned Behavior serve as an appropriate theoretical model to

measuring teachers' IEP data collection intentions, such that teachers' attitudes, subjective norms, and perceived behavioral control explain a significant amount of the variance in ECSE teachers' intent to engage in future IEP data collection?

Dissertation Organization

This dissertation is organized into four subsequent chapters. Chapter Two includes a broad overview of the various bodies of literature related to the exploration of IEP data collection practices in ECSE as well as the creation of future implementation supports geared toward reducing the research-to-practice gap. Chapter Three contains a comprehensive description of all study procedures, from scale creation to data analysis, and Chapter Four presents the results associated with each analysis conducted. In conclusion, an in-depth discussion of the results, including all related study limitations and future directions is included in Chapter Five.

Chapter Two – Literature Review

The frequent collection and use of data for the purposes of child-specific instructional planning is not only legally required of special educators (IDEA, 2004), but it has also been associated with higher quality instruction and improved outcomes for students receiving special education services (e.g., Fuchs et al., 1984; Mirkin, et al., 1982; Stecker & Fuchs, 2000). In order to ensure program effectiveness and student progress, it is imperative that teachers engage in the ongoing collection of data highlighting students' performance of skills targeted in their Individualized Education Programs (IEPs). Regardless of the legal requirements and empirical evidence, however, a research-to-practice gap persists when it comes to data collection (Brawley & Stormont, 2013; Cook et al., 1991; Fuchs, Fuchs, & Warren, 1982; Sandall, Schwartz, & LaCroix, 2004; Wesson, King, & Deno; 1984).

Given that "[ECSE] services represent an evolving professional specialization that is qualitatively different from the special education services provided to school-age populations especially in regard to the planning and delivery of individualized services" (Garrett & Kelley, 2012, p. 269), it is plausible to expect that ECSE teachers experience unique challenges related to IEP data collection; challenges that might never be exposed in the K-12 literature. Regardless of the potentially distinctive adversities faced by ECSE teachers when monitoring their students' progress, however, data collection is a necessary part of the profession. As a requirement for initial licensure, ECSE teachers must demonstrate their ability to select, adapt/modify, administer, and interpret assessment measures for young children for the purpose of evaluating, planning, implementing, and monitoring the IEP (Minnesota Rule, 8710.5500, 2017). Additionally, included in the Council for Exceptional Children's Division for Early Childhood's (DEC) recommended practices, ECSE teachers are encouraged to implement assessment practices that are both systematic and ongoing in order to plan instruction that is responsive to children's progress (2014). To aid in the development of effective implementation supports allowing ECSE teachers to uphold their professional responsibilities, increased attention toward IEP data collection practices in ECSE is crucial.

Before examining the literature specific to data collection in ECSE, existing frameworks, procedures, and tools created for and available to educators that support their data collection practices will be explored. Approaches specific to early childhood (EC) will be highlighted and the appropriateness and feasibility of these approaches in monitoring the progress of young children with special needs will be considered. Following the review of data collection facilitators, the literature specific to ECSE professionals' data collection practices will be examined in hopes to better understand the various implementation barriers and study limitations possibly adding to the research-to-practice gap in this area. Next, theoretical frameworks highlighting the important role played by a key individual-level factor—motivation—when attempting to facilitate teachers' behavior change will be reviewed.

Finally, the Theory of Planned Behavior (TPB; Ajzen, 1985), which is hypothesized to support the future exploration of the many factors impacting ECSE teachers' intentions to engage in data collection will be presented and discussed in detail. A recent application of the TPB in better understanding and predicting special educators' IEP data collection practices will be reviewed, with a focus on how each construct was defined and measured. In order to support the creation, evaluation, and modification of implementation supports aimed to increase ECSE professionals' collection of IEP data, recommendations for the future application of enhanced measurement techniques will be presented.

Facilitators of Data Collection

Knowing that progress monitoring is a critical process in the education of young children leading to important decisions that improve the individualized instruction provided to students (DEC, 2007), a variety of supports—from laws and structured frameworks to individual assessment tools and various technologies—have been developed to facilitate the collection of data within the progress monitoring cycle. Following is a brief description of these supports, including references to the evidence highlighting their impact on teacher's data collection practices when available.

Individualized Education Programs. Essential to Public Law 94-142 (P.L. 94-142; 1975), the Individual Education Program (IEP) was created as both a process and a document to ensure students received the appropriate education they were legally entitled to (Smith, 1990). Defined by Christle and Yell (2010) as substantive requirements, the *IEP document*, hereafter referred to as the IEP, is required to include measurable goals linked to educational needs and a description of how the educational team plans to measure educational progress (IDEA, 2004). Though the IEP contains numerous other mandatory components, it is the division of goals into specific and measurable objectives and the plan for tracking progress that helps to facilitate educators' collection of data. High quality objectives and detailed plans for measuring progress should leave little to no questions when it comes to determining what skill to observe, under what conditions the skill needs to be observed, how often the observation needs to occur, and what type of data is to be collected.

Multi-Tiered Systems of Support. Designed as a framework to support the provision of individualized and evidence-based instruction that is grounded in the assessment of student progress, multi-tiered systems of support (MTSS) has been widely used in K-12 education settings to address both the academic and social, emotional, and behavioral needs of students. Response to Intervention (RTI), an application of MTSS, provides teachers with a structured framework for assessing student's academic progress and making data-based instructional decisions, leading to the implementation of evidence-based practices individualized to the needs of all students (Fuchs & Fuchs, 2006). As opposed to highly structured protocols used by researchers, most school systems employ a flexible problem-solving approach to guide the collection and use of data (Fuchs & Fuchs, 2006). This approach to problem solving not only helps teachers to answer questions regarding what tools should be used to evaluate student progress and how often these tools should be used, but it also assists teachers in their use of data by suggesting patterns of student performance that indicate a need for a modified instructional plan as well as recommending instructional changes.

Though less commonly instituted in early childhood settings, the Center for Response to Intervention in Early Childhood (CRTIEC) is dedicated to conducting research and creating resources that support the implementation of RTI in early childhood settings (e.g., Greenwood et al., 2015; McConnell, Wackerle-Hollman, Roloff, & Rodriguez, 2015). While the impact of RTI on teacher's data collection practices in ECSE is unknown, there is evidence to suggest that this type of structured framework has the potential to enhance teachers' decision making (VanDerHeyden et al., 2007) and impact child outcomes (Gettinger & Stoiber, 2007).

In addition to RTI, Positive Behavioral Interventions and Supports (PBIS) is another application of MTSS that supports teachers' data collection. Directly targeting students' social, emotional, and behavioral functioning, thus subsequently impacting their academic success, PBIS is typically implemented schoolwide and often consists of three levels of support: 1) primary prevention directed toward all students, 2) secondary prevention directed toward a relatively smaller group of students needing more concentrated support, and 3) intensive intervention directed toward the smallest group of students in the form of functional behavior assessments and behavior intervention plans (Sugai & Horner, 2006). The Pyramid Model, an application of schoolwide PBIS in early childhood, provides teachers with not only a framework for the collection and use of data related to young children's social, emotional, and behavioral development, but also a specific tool—the Behavior Incident Recording System (Fox, Perez Binder, Liso, & Duda, 2010)—to monitor behavior incidents across programs (Fox, Veguilla, & Perez Binder, 2014).

Curriculum-Based Assessments. The most commonly used method of data collection to monitor student progress is the use of curriculum-based assessments (CBAs). Researched extensively, CBAs are a specific type of assessment that a) include familiar testing materials drawn from the student's curriculum, b) occur repeatedly over time, and c) are used to evaluate student performance in order to make responsive instructional decisions (Fuchs & Deno, 1991). According to Fuchs and Deno (1991), most forms of CBAs can be placed into one of two categories—*specific subskill mastery measurement* or *general outcome measurement*.

Specific subskill mastery measurement, also known as a critical mastery approach, allows teachers to collect data on sequences of discrete skills that lead up to broader developmental outcomes (McConnell, 2000). In early childhood, numerous CBAs employing the subskill mastery approach to assessment are available for teachers to use to monitor child progress including but not limited to the Assessment, Evaluation, and Programming System for Infants and Children (AEPS; Bricker, Cripe, & Slentz, 2003); Hawaii Early Learning Profile (HELP; VORT Corporation, 2010); Teaching Strategies: GOLD (TS: GOLD; Teaching Strategies, Inc., 2011); and Work Sampling System (WSS; Meisels, Jablon, Marsden, Dichtelmiller, & Dorfman, 2001). While little is known about the impact of CBAs on teachers' collection of IEP data in early childhood settings, there is some evidence to suggest that use of the AEPS is linked to the creation of higher quality IEP goals and objectives (Pretti-Frontczak & Bricker, 2000), which is imperative to subsequent data collection.

General outcome measurement, on the other hand, focuses on repeated measurement of the broader outcome a child is working toward, allowing for the repeated measurement of a skill over time using consistent data collection methods, regardless of how the skill is task-analyzed or what instructional method is used (Fuchs & Deno, 1991; McConnell, 2000). Though less commonly used in early childhood, several assessments specific to monitoring progress in young children utilizing a general outcome measurement system exist including the Individual Growth and Development Indicators (IGDIs) for Infants and Toddlers (Greenwood et al. 2011b; Missall et al. 2008) and the

Preschool IGDIs (Greenwood et al., 2006, Greenwood et al., 2001b; Roseth et al. 2012). While the utility of these tools in monitoring young children's progress toward achieving broad developmental goals is evident (e.g., Greenwood, Carta, & McConnell, 2011), the extent to which these tools support teachers' collection of IEP data in ECSE is unknown.

Technology. In addition to IEPs, structured frameworks for collecting and analyzing data, and numerous data collection tools, various computer-based decision support systems (CDSSs) have been created to support the collection and use of data across developmental domains in early childhood (Buzhardt et al., 2012). Making Online Decisions (MOD; Buzhardt et al., 2010), a component of the Infant and Toddler IGDI online data system, guides teachers through a decision-making process resulting in intervention recommendations, all grounded in child-specific data. Use of MOD by early childhood professionals in a home visiting program has been found to significantly increase young children's expressive language skills when compared to children whose home visitors did not receive MOD support (Buzhardt, Greenwood, Walker, Anderson, Howard, & Carta, 2011).

Another technology worth noting is the Knowledge Management System for Behavioral Incidents (KMS-BI; Johnson, 2017). Created in response to early childhood professionals' need to easily collect, interpret, and share data regarding children's behavior; this example of a cloud-based problem solving system allows teachers to electronically document incidents of challenging behavior and access automatically generated visual reports summarizing behavioral data for children, classrooms, and/or programs. Finally, though not created specifically for use in ECSE, the DDTrac software allows teachers to electronically monitor their students' progress toward accomplishing

IEP goals by individualizing the type of data (i.e., current/incorrect, correct/incorrect with prompt, no response, rating scales, scores, duration, and frequency) they plan to collect across objectives and visually summarizing the data. Each of these technologies aims to increase the feasibility and utility of teachers' data collection and use.

Barriers to Data Collection in Early Childhood Special Education

Regardless of the availability of supports that have the potential to facilitate the frequent collection of data for child-specific instructional planning in ECSE, a number of subjective data collection methods continue to persist. In addition to trusting personal memories in the absence of recorded data (Sandall et al., 2004), anecdotal notes taken during informal observations make up a large portion of early childhood educators' documentation of child progress (Brawley & Stormont, 2013). When attempting to understand the challenges faced by ECSE professionals when it comes to the systematic collection of data, it is crucial to attend not only to enhancing educator skills, but also to understanding their beliefs and subsequent motivation to engage in data collection as well as the context in which they work, as these factors impact the degree to which the practices are accepted and implemented (Aarons, Hurlburt, & McCue Horwitz, 2011; Han & Weiss, 2005; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Michie, van Stralen, & West, 2011; Sparks 1988).

Following is a comprehensive review of the literature highlighting factors hypothesized to serve as barriers to the frequent collection of data for the purposes of child-specific instructional planning in ECSE. In an attempt to add clarity to the following discussion, potential factors will be organized into two categories—internal and external. Internal factors are thought to include variables characteristic of a specific individual, while external factors are variables characteristic of the context in which the individual works. Though described independently, many internal and external factors are theorized to be associated, as highlighted by the overlapping circles in Figure 1.

Internal Factors Impacting Data Collection. As highlighted by Fixsen et al. (2005), ineffective programs can be implemented well, and effective programs can be implemented poorly—both of which are problematic. It is important, therefore, for both a teacher's knowledge and skill to be considered in relation to data collection. In addition to one's capacity for data collection, additional variables—more subjective in nature—thought to impact data collection behaviors include an individual's perception of their capacity (i.e., self-efficacy), in addition to their motivation to collect data. Just as an organization's motivation to engage in evaluation has been shown to impact the organization's capacity to evaluate (Suarez-Balcazar et al., 2010), an individual's motivation to collect data likely impacts their future data collection practices. Factors thought to impact a teacher's motivation include their perceptions of the importance of the IEP and subsequent data collection, their perceptions of the practicality of ongoing data collection, and their perceptions of the demand to modify their existing data collection practices.

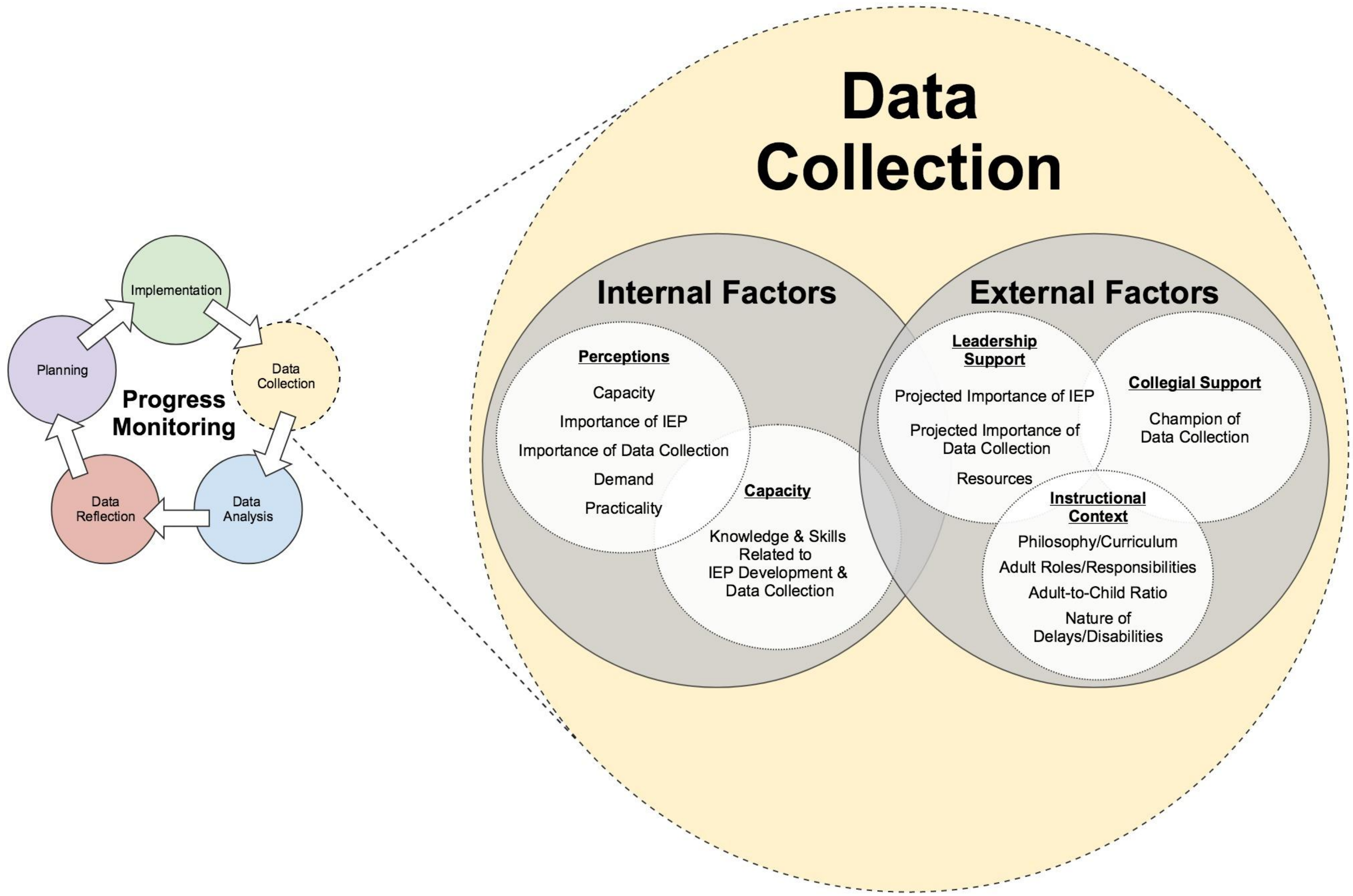

Figure 1. Factors Impacting Data Collection. This figure illustrates the internal and external factors hypothesized to influence teachers' data collection practices in ECSE.

Capacity. An individual's capacity to collect data refers to their *knowledge* and *skill* related to IEP development and subsequent data collection. Knowledge is what a teacher knows based on pre-service and in-service training activities and is influenced by their self-reflection on past experiences, while skill refers to the teacher's actual ability to write clear and measurable IEP objectives and collect data on student progress toward meeting each IEP objective. While the literature on teachers' actual capacity (versus self-reported knowledge and skills) related to data collection is nonexistent, it is hypothesized that both a lack of knowledge (i.e., skill deficit) and a lack of actual skill (i.e., performance deficit) will serve as a barrier to data collection.

Perceived Capacity. Measured more frequently than one's actual capacity due to the nature of previous investigations (i.e., qualitative methods consisting primarily of surveys), an individual's perception of their capacity is also thought to impact their data collection practices. Perceived capacity includes an individual's beliefs about their knowledge and skills related to data collection. Consistent with Bandura's notion of self-efficacy (1971), a teacher may believe data collection leads to more effective teaching and improved child outcomes, but if they do not perceive themselves as possessing the knowledge and skill needed to effectively engage in the frequent collection of IEP data, it is hypothesized that the teacher will put forth less effort to collect IEP data and is likely to identify more barriers to data collection. Teachers who are confident in their capacity, however, are hypothesized to be better equipped to overcome common barriers to data collection, thus reducing the effort required of future data collection. A teacher's decreased effort, increased data collection practices, and reinforcement in the form of effective instructional planning with ease and improved student outcomes, creates a motivating cycle potentially decreasing the number of reported barriers.

Based on the results of a qualitative study examining ECSE professionals' perspectives on data collection (Sandall et al., 2004), participants reported that lack of knowledge and issues with data management contributed to their sporadic and unstructured data collection practices. Many participants noted that they often collected more frequent data on the behaviors of certain children, leaving other children with little to no data highlighting their performance. Additionally, teachers stated that structured interventions (e.g., toilet training programs, medication trials) tend to facilitate their data collection, while the collection of data during instruction that is embedded into naturally occurring classroom routines was more difficult. This inconsistency in teachers' perceived capacity across skills and interventions potentially indicates a lack of training and/or experience regarding various data collection methods. Furthermore, though not explicitly linked to their data collection capacity, participants reported feeling dissatisfied with the data collection procedures they were using, which could also be attributed to limited knowledge or lack of skill.

Concerning the data management issues, participants reported having a difficult time determining and assigning data collection roles, collecting data on multiple objectives across multiple children, and communicating information pertaining to data collection to all classroom staff (Sandall et al., 2004). Cooke et al. (1991) also highlighted a possible lack of knowledge and/or skill related to the establishment of data management systems, based on the number of teachers (97.6%) who reported collecting all student data themselves; only 14.5% of surveyed teachers reported having a classroom aide help in the collection of data. Again, while not explicitly stated, these issues could possibly be due to lack of skill or limited knowledge of data management strategies.

In addition to the skill-related barriers to data collection overtly described by Sandall et al. (2004), participants also noted that the nature of their student's IEPs impacted their data collection practices. ECSE professionals reported that many IEPs did not reflect behaviors viewed as important targets in the classroom. Additionally, it was reported that student progress observed in the classroom was often not reflected in the IEPs. Though not described to be directly related to educator skill, it can be implied based on previous research (Notari & Bricker, 1990) that this phenomenon was related to the participants' ability to effectively monitor student progress and link resulting data to the creation of functional, specific, and measurable IEP objectives, sensitive to each student's rate of progress.

Perception of Importance of IEP. Following the passage of the P.L. 94-142 in 1975, which prompted the IEP mandate, research on teachers' perceptions of the IEP revealed a common perspective that the IEP was more a legal obligation than an important document guiding instruction (Dudley-Marlin, 1985; Margolis & Truesdell, 1987; Morgan & Rhode, 1983). Teachers reported that the IEP had little to no bearing on the quality of their instruction or on their students' learning (Dudley-Marlin, 1985; Margolis & Truesdell, 1987; Morgan & Rhode, 1983), generating a shared opinion that IEPs were unnecessary to the provision of individualized instruction. Because IEP development initiates the progress monitoring cycle, and within the cycle, data is collected on students' progress toward meeting IEP objectives, the connection between teachers' perceived importance of IEPs and their data collection practices is clear.

Though no data exist to support (or dispute) this claim, it is hypothesized that teachers who believe IEPs are fundamental to both quality instruction and student growth, will be more motivated to engage in frequent data collection practices. Therefore, an underlying belief that IEP development is a meaningless waste of time serves as a potential barrier to the collection of IEP data.

Perception of Importance of Data Collection. Teachers are more motivated to adopt a practice if they find it important (Sparks, 1988) and believe it will benefit their students (Klingner et al., 2003). This is encouraging, as the link between frequent measurement of students' skills and improved student outcomes has been clearly established. There is evidence to suggest, however, that due to a lack of trust in educational research, some teachers chose to rely on their own experiences whether or not they are in alignment with what the research suggests (Boardman et al., 2005; Cook & Cook, 2011; Foster, 2014; Hornby, Gable, & Evans, 2013). It is hypothesized, therefore, that an individual's perception of the importance of data collection potentially serves as a barrier to the frequent collection of child-specific data in ECSE.

Based on the available survey data, most ECSE professionals reported that data collection was important (Browley & Stormont, 2013; Ruble et al., 2018; Sandall et al., 2004). Many of these same educators, however, seemed confident in their ability to accurately judge a child's progress in the absence of data and trusted use of their own memories of child performance when measuring progress (Sandall et al., 2004). The reason for believing the data collected are not meaningful or do not accurately represent children's skills remains unclear; however, it is hypothesized to be related to their current data collection practices (e.g., those who engage in frequent data collection are more apt to believe the data hold meaning, while those who rely on subjective observations or memories might lack belief in the importance of the data), their satisfaction with these practices, and their capacity to collect data.

Perception of Demand. Individuals are more motivated to adopt a practice if there is a need for change (e.g., Fixsen et al., 2005); therefore, a teacher's perception of the demand for data collection is hypothesized to impact subsequent data collection. For teachers who are already engaging in data collection practices, the degree to which they are satisfied with their current practices is thought to be inversely related to their perceived demand for a new or slightly modified method of data collection. For example, a highly satisfied teacher, regardless of the quality of their practices, is less apt to adopt evidence-based data collection practices, as there is no perceived demand to do so. To date, Sandall and colleagues (2004) are the only researchers to explore ECSE professionals' satisfaction with their own IEP data collection practices. In this study, most teachers reported being dissatisfied with their current practices and all reported a desire to make a change. Though not directly linked to satisfaction, a number of empirical examinations and literature reviews have identified challenges faced by ECSE professionals when attempting to collect data (Akers et al., 2014; Banerjee & Luckner, 2013; Cooke et al., 1991; Ruble et al., 2018; Sandall et al., 2004; Zweig et al., 2015), which potentially impacts their satisfaction with data collection. Based on these data, it is hypothesized that a perceived demand likely influences subsequent data collection practices.

Perception of Practicality. In addition to a teacher's perception of demand, the perceived practicality of data collection practices is hypothesized to influence their motivation to engage in data collection. One facet of practicality highlighted by Doyle & Ponder (1977) is *congruence*, which is the degree to which the practice aligns with current practices. Though not yet explored in relation to data collection in ECSE, we know that practices that are more closely aligned with a teacher's current instructional practices (whether chosen by the individual or required by administration) are more likely to be implemented (e.g., Klingner et al., 2003; Sparks, 1988). *Ease of use* is another component of practicality as described by Doyle & Ponder (1977). Again, while little is known about ECSE teachers' perceptions of the practicality of the frequent collection of data for child-specific instructional planning, we know that how easy or challenging a practice is perceived has been shown to influence implementation (e.g., Aarons, 2004).

External Factors Impacting Data Collection. In addition to the many internal factors characteristic of individual teachers, the literature suggests that the environment within which a teacher works can present a range of barriers –school systems, class size, available resources, etc.—that will likely impact their ability to implement evidence-based practices (Foster, 2014). The external factors distinctive of the diverse contexts in which ECSE teachers work that are hypothesized to influence their data collection practices include leadership support, collegial support, and instructional context.

Leadership Support. Leadership support is imperative to the facilitation of implementation (e.g., Fixsen et al., 2005; Klingner et al., 2003). An administrator's belief that quality IEPs and the frequent collection of data to guide instruction is important as well as the availability of resources such as time, materials, and professional development is thought to facilitate teachers' use of a practice (e.g., Banerjee & Luckner, 2013; Sandall et al., 2004; Taylor, Nelson, & Adelman, 1999). Limited leadership support, therefore, is hypothesized to serve as a barrier to teachers' collection of data in ECSE.

Projected importance of IEP and data collection. Publicity of a leader's belief in and commitment to a practice is thought to facilitate teachers' use of the practice (e.g., Taylor et al., 1999). Little is known, however, about the perceptions of those in ECSE leadership roles regarding the importance of the IEP and the ongoing collection of child-specific data used to monitor progress and plan instruction. Similarly, the literature describing how these individuals' project their beliefs regarding their importance onto the teachers they support is scarce. In regard to leaderships' projected beliefs about the importance of data collection, Ruble et al. (2018) found that most ECSE teachers reportedly viewed their administration as believing in the importance of data collection and believed those in administrative roles would support their data collection practices. These same teachers, however, were only slightly more likely to agree than disagree that these same individuals pay attention to the data they collect. Similarly, Brawley and Stormont (2013) reported that 33% of early childhood educators stated their administrator was rarely or never involved in the analysis of student data. These data shed light on a possible inconsistency between how teachers view others to believe in and support their data collection and how others attend to and use the data.

Though not reflected in the literature, increased state accountability may potentially be feeding this inconsistency. ECSE professionals are held accountable for measuring and reporting on student progress in order to demonstrate program effectiveness (Kasprzak et al., 2012); however, the type of data they are required to report is related to general early childhood outcomes, rather than individual educational goals.

Despite the lack of evidence supporting this hypothesis, it is possible that teachers are getting the message (whether directly or indirectly) that infrequent data collection for state accountability purposes is more valued than ongoing, individualized progress monitoring; the latter of which holds professionals accountable to the children and families they serve.

Furthermore, keeping in mind that data collection is driven by the creation of IEPs, it is imperative to explore how teachers perceive to be pressured by administrators to write high quality IEPs that include specific and measurable objectives. Based on reports of inconsistencies between skills targeted in IEPs and skills viewed as important to everyday functioning in the classroom, Sandall et al. (2004) suggested a possible need for increased administrative support to facilitate more frequent revisions of IEPs resulting in greater quality. This, however, is assuming those in leadership roles view the IEPs as not only a due process requirement, but also as a guide to individualized instruction and data collection.

Resources. In order to engage in frequent data collection practices, it is imperative that teachers are provided adequate time, materials, and professional development. Based on teacher report, the main contextual factor impacting the data collection practices of ECSE professionals is lack of time (Banjeree & Luckner, 2013; Cooke et al., 1991; Sandall et al. 2004). Teachers reported not having enough time to collect, analyze, interpret, and share data. Additionally, they reported that data collection practices seemed to take valuable time away from preparing for or providing instruction. Teachers' time allocation in ECSE, however, has not been analyzed; therefore, it is unclear whether there is truly a lack of time—a barrier that could potentially be eased by administrators—or whether teachers are inadvertently distributing their time inefficiently due to their level of capacity and/or their perceptions of importance regarding data collection.

Conflicting with earlier explorations of barriers to data collection, a more recent study suggests that variables related to time—too little time, too much to do, and too many students—generally do not impact teachers' data collection (Ruble et al., 2018). Teachers surveyed as part of this study, however, agreed that "unclear measurement systems" do impact their ability to collect data in a timely matter. Similar to this finding, Banerjee and Luckner (2013) and Sandall et al. (2004) found that teachers commonly identified the lack of appropriate tools as a barrier to data collection. In addition to time and materials, Sandall et al. (2004) identified a need for continuous and individualized support in order for ECSE teachers to embed data collection practices into existing routines as well as to summarize, analyze, and interpret the data for the purposes on instructional planning. Other examinations related to progress monitoring in ECSE have made connections between teachers' capacity and their subsequent data collection practices (Cooke at al., 1991; Notari & Bricker, 1990; Sandall et al., 2004), further suggesting a need for access to professional development activities. Additionally, performance feedback—a form of professional development—has been described to increase teachers' perceptions of demand, as the feedback provided can highlight connections between a teacher's effective practice and various positive changes in the classroom (Han & Weiss, 2005). Though not overtly linked to leadership support, lack of materials used in the collection of data and the need for additional training could potentially be addressed by administrators.

Collegial Support. In addition to leadership support, the presence of a collegial advocate of any given practice might impact a teacher's use of that practice. An individual or group of individuals dedicated to implementing a practice with fidelity and promoting the efficacy of the practice among colleagues—described by Fixsen et al. (2005) as "purveyors"—can facilitate widespread use of the practice across an organization. Though not addressed in the literature, it is hypothesized that limited collegial support serves as a potential barrier to one's data collection in ECSE.

Instructional Context. According to recent child count data, only 38% of children between the ages of 3 and 5 who are eligible for special education services, receive services in an inclusive early childhood program (USED, 2015). Twenty-three percent receive services in a separate classroom (e.g., special education only classroom) and the remaining 34% receive services in some other location (e.g., home, residential facility, separate school). With this range of settings, comes an even wider range of instructional contexts, each thought to have a unique influence on teachers' data collection practices. While impacted by setting, instructional contexts are thought to be influenced more heavily by teaching philosophy, curriculum, teacher roles and responsibilities, adult-to-child ratio, and the nature of children's delays/disabilities. For example, one inclusive preschool classroom might follow a structured curriculum comprised of numerous teacher-directed activities, while the activities in another inclusive preschool classroom are mostly child-directed, grounded in a program philosophy, but without use of a specific curriculum. Though the setting is the same (i.e. inclusive early childhood program) across both classrooms, the instructional contexts created by the teaching philosophy and curriculum are very different.

In addition to diverse philosophies and curricula across classrooms, an ECSE teacher's roles and responsibilities vary greatly across instructional contexts, which likely impacts their data collection practices. In a separate special education classroom attended only by students with IEPs, the ECSE teacher—along with other ECSE service providers and support staff—is likely responsible for all of the instructional planning and delivery. In an inclusive setting, however, the roles and responsibilities of the ECSE teacher fluctuates greatly from one classroom to the next; from minimal planning and delivery of instruction or sharing responsibilities with the general education teacher, to full responsibility for all teaching-related activities (i.e., reverse mainstreaming). Along with teacher roles and responsibilities, the number of teachers in relation to the number of children (i.e., adult-to-child ratio) in any given instructional context is hypothesized to serve as a potential barrier to data collection.

Finally, the nature of the delays and disabilities represented in and across classrooms creates another challenge unique to ECSE teachers when it comes to data collection. The heterogeneous needs of children receiving ECSE services creates a situation in which teachers must be prepared to engage in progress monitoring practices that align with a diverse set of educational goals and interventions. In alignment with Deno and Mirkin's statement that "no method or system of individualized programming is adequate for all children, in all classroom settings, and under all circumstances" (1977, p. 5), it is likely that ECSE teachers need to simultaneously employ various methods of data collection in order to individualize their progress monitoring practices to meet the needs of each child, perhaps increasing the amount of knowledge/skill, time, materials, and support that is necessary to collect data. Additionally, given the developmental level and rate of this particular group of children, they are often working toward very specific, functional skills not often highlighted in commercially available assessments. As Downs and Strand (2006) described, while data from CBMs (and other general outcome measures) offer evidence of a child's progress over time in relation to general early childhood outcomes, they may not be sensitive enough to inform the instruction of many preschoolers with special needs. Because of this, many ECSE teachers are left to create their own data collection tools (Banerjee & Luckner, 2013; Brawley & Stormont, 2013) or modify existing tools to better meet their needs. Additionally, many teachers continue to record anecdotal notes as a method of monitoring student progress (Brawley & Stormont, 2013), which is arguably more subjective and time consuming than other methods.

Just as ECSE teachers are expected to individualize their data collection procedures to meet the extensive range of needs inherent in their students (Odom & Wolery, 2000), they must also match their methods to the environment in which they work. While the influence of setting on ECSE teachers' data collection practices has yet to be systematically explored, there is some evidence to suggest that instructional context potentially serves as a barrier to the frequent collection of data.

Limitations of the Literature Base

The many challenges faced by ECSE teachers when attempting to engage in the frequent collection of data for the purposes of child-specific instructional planning is evident. Though a wide range of barriers has been identified, the size of the literature base is negligible; only six studies have been explicitly described to include the examination of teachers' data collection practices in ECSE (Banerjee & Luckner, 2013;

Browley & Stormont, 2013; Cooke et al., 1991; Luckner & Bowen, 2006; Ruble et al., 2018; Sandall et al., 2004). Of these studies, only three are exclusive to the data collection practices of early childhood educators (Banerjee & Luckner, 2013; Browley & Stormont, 2013; Sandall et al., 2004), while only one focused solely on the data collection of ECSE professionals. Knowing the important role data collection plays in the progress monitoring cycle and keeping in mind the previously discussed barriers potentially unique to the collection of IEP data in ECSE, continued research in this area is necessary to the creation of effective training programs. Before moving forward, however, it is important to briefly consider some of the limitations in the literature in order to strengthen future research in this area. In addition to lacking an organizing framework, the literature lacks clearly defined constructs; both of which if addressed, are hypothesized to facilitate the creation of appropriate supports aimed at closing the research-to-practice gap.

Lack of Organizing Framework. Along with the size of the literature base, the absence of a guiding framework grounded in theory is perceived as a current limitation. This type of framework has the potential to not only aid in the identification of important variables, but also to guide the measurement of such variables as well as subsequent data analysis. Potentially adding to the lack of theory-driven research in this area is the diverse set of views represented in the field of psychology, resulting in numerous theories generated to help explain behavior. Behaviorism, an objective branch of psychology involving the study of observable behaviors and their relationship with environmental stimuli that set the occasion for and maintain behavior (Watson, 1913), is supported by a vast amount of empirical evidence dating back to the turn of the 20th century (Pierce &

Cheney, 2013). While the importance of behaviorism is hard to overlook, it is only one of many perspectives on behavior and learning. Cognitive psychology, another approach to explaining behavior, involves the direct study of covert behaviors such as thinking and feeling (David, Miclea, & Opre, 2004). While strict behaviorists are only concerned with the analysis of observable behaviors, contemporary behavior analysts include internal events such as thoughts and feelings as part of the environmental context in which behaviors occur (Pierce & Cheney, 2013). Those with a cognitive perspective, on the other hand, argue that observable behaviors can be explained by an individual's thoughts and feelings.

While these two approaches to explaining behavior are considerably different from one another, both theoretical orientations have grown closer over time (Dowd, Clen, & Arnold, 2010). This blurring of lines between behaviorism and cognitive psychology is especially beneficial when studying the behaviors of teachers. Recently, Ruble and colleagues (2018) were the first to use the Theory of Planned Behavior (TPB; Ajzen, 1985)—defined as a social-cognitive theory highlighting the connection between beliefs and behavior—to guide their examination of the variables that facilitate and hinder the data collection practices of special educators of young children with autism. In doing so, they were able to gain a preliminary understanding of how teachers' beliefs and intentions support their data collection behaviors.

Lack of Clearly Defined Constructs. In addition to the lack of an organizing framework, the literature is inundated with partially and inconsistently defined constructs, creating a challenge when it comes to interpreting the results of individual studies as well as comparing results across studies. Almost exclusively qualitative in nature, every study represented in the literature base included the administration of a survey, resulting in a general consensus that ECSE teachers "believe" data collection to be an important practice (Brawley & Stormont, 2013; Cooke et al., 1991, Ruble et al., 2018; Sandall et al., 2004). Confusion continues to exist, however, around what constitutes a belief as well as how beliefs should be measured. Described by Pajares as "travel[ing] in disguise and often under alias", beliefs have been referred to by many names in the literature, including but not limited to attitudes, values, opinions, preconceptions, perceptions, perspectives, personal theories, and rules of practice (1992, p. 309).

Though not operationally defined in this way, the term *belief* in the current literature appears to represent teachers' self-reports of the *importance* of data collection. In a survey of special educators (including, but not limited to those in early childhood), Cooke et al. (1991) asked participants about the *importance* of data collection in determining student progress, yet later referred to these data as representing the teachers' *beliefs*. Sandall et al. (2004) were reportedly interested in collecting data on ECSE professionals' *perspectives* on data collection and use, which were gathered via a combination of survey, interview, and reflective writings. In summarizing the results, Sandall et al. (2004) organized information regarding the *importance of*, *demand for*, and *satisfaction with* data collection—variables often encompassed in the assessment of motivation—as well as *current practices* under a category labeled "beliefs and practices". In alignment with previous studies, Brawley and Stormont (2013) collected data on early childhood professionals' ratings of *importance* across a variety of data collection procedures, which were later described as their *beliefs*. While an organizing framework grounded in theory will likely lead to more thorough and consistent defining of constructs, operationally defined constructs are necessary regardless of such frameworks.

Similar to the vague descriptions of beliefs present in the literature, the behavior of interest here—data collection—lacks a clear and consistent definition. More broadly speaking, Banerjee and Luckner (2013) attempted to gain a better understanding of the "assessment practices" of early childhood educators. Included in their survey was a common definition of assessment, which included methods used for screening, diagnosis, instructional planning, placement, progress monitoring, and program evaluation. Though progress monitoring was included, it was not defined, leaving it open to individual interpretation. Similarly, Luckner and Bowen (2006) surveyed teachers of the deaf and hard of hearing (including those in early childhood) regarding their current assessment practices. Respondents were provided with a definition of assessment—"a process of collecting data for the purpose of making decisions about individuals and groups"—but given the generality of the definition, the data were likely construed knowing the various ways in which teachers could have been thinking about data and decision making (p. 412). As opposed to assessment practices, Brawley & Stormont (2013) examined early childhood educators' perceptions of the importance of various *data collection practices* as well as their frequency of use. While each survey question focused on a specific method and/or purpose of data collection, items lacked definitions of "goals" and "progress monitoring". In each of these investigations, while it can be assumed that the resulting data represent, at least in part, teachers' views and practices related to the collection of data highlighting progress toward IEP goals, this assumption cannot be proven.

More closely aligned with the variable of interest described in this paper—the frequent collection of data representing children's progress toward meeting IEP goals for the purposes of planning individualized instruction—Cooke et al. (1991), Sandall et al. (2004), and Ruble et al. (2018) were all interested in studying the collection of data specific to IEP goals. Cooke et al. (1991) not only titled a section of their survey, "Measuring Student Progress Toward IEP Objectives" (p. 156), but they also included the phrases "objective data" or "IEP objectives" in many of their survey items. Sandall et al. (2004) identified three types of data they wanted ECSE professionals to think about when answering interview and survey questions, one of which included "weekly documentation for each [IEP] objective" (p. 164). Additionally, participants were specifically asked how these child data were used in lesson planning.

Finally, Ruble et al. (2018) surveyed teachers of students ages 3 to 8 years with autism to better understand the teachers' intent to collect data, self-efficacy related to data collection, and views of administrative support related to their collection of data, as well as their actual data collection practices. While they explicitly stated they were interested in "teachers' views of data collection for IEP goals", only some of their survey items explicitly linked data collection to IEP goals (Ruble et al., 2018, p. 4). Furthermore, they measured teachers' actual ability to collect data on skills targeted as part of a separate intervention study, but it was not clear if and how the targeted skills were related to the students' IEP goals. Keeping in mind the many purposes driving data collection in ECSE—eligibility determination, instructional planning, and state accountability to name a few—it is imperative that "data collection" be operationally defined in future research in order to facilitate the validity of measurement approaches as well as the accuracy of interpretations based on the data.

The identification and defining of constructs using clear and mutually agreeable terminology has been recognized as a critical first step in the measurement process as well as a common measurement problem (Thorndike & Thorndike-Christ, 2010). Specific to the measurement of constructs via surveys, the predominant method used in this literature base, Dillman, Smyth, and Christian's (2014) instruct survey designers to "use specific and concrete words to specify the concepts clearly," so as not to leave room for interpretations by respondents (2014, p. 119). Moving forward, it will be important for researchers to attend to these guidelines when constructing measures and reporting results.

Theoretical Guidance for Addressing the Research-to-Practice Gap in ECSE

Although limitations exist and questions remain, previous research has uncovered many details about data collection practices in ECSE that will be useful when attempting to reduce the gap between research and practice. We are aware of the many benefits of progress monitoring (Fuchs et al., 1984; Fuchs et al., 1989; Mirkin et al., 1982; Stecker & Fuchs, 2000) and we know that ECSE teachers generally report data collection to be an important practice (Brawley & Stormont, 2013; Cooke et al., 1991; Ruble et al., 2018; Sandall, et al., 2004). We also know that while it is perceived to be important, the frequent collection of data regarding IEP goals continues to be underutilized by ECSE professionals (Brawley & Stormont, 2013; Sandall et al., 2004). Numerous facilitators and barriers of data collection have been identified that help explain this gap between research, reported beliefs, and reported practices; however, it is not clear how these factors are related due to the organization and clarity of the literature. Furthermore, the degree to which teachers' beliefs and reported practices are truly discrepant, and whether or not this gap has been influenced by unsatisfactory measurement is unclear. To assist in the creation of training programs that better prepare and support ECSE teachers in the area of data collection, a practice that is compatible with any and all instructional contexts, it is crucial that future research address these limitations.

While researchers in implementation science have been successful in identifying distinct characteristics of training programs, administrators, coaches, and practitioners that lead to implementation of various practices with greater fidelity, less is clear regarding the impact of inherent characteristics on teachers' adoption or implementation of practices (e.g., Odom, 2009; Fixsen et al., 2005). In working toward the implementation of evidence-based practices in the public health domain, Michie et al. (2011) based their analysis of behavior change interventions and subsequent creation of a framework to guide future behavior change interventions on the COM-B system. The COM-B system—which places an emphasis on individual-level factors—highlights the importance of considering a person's *motivation* to change their behavior, which is impacted by their *capability* to engage in the new behavior as well as the *opportunity* to engage in the new behavior. Together, one's capability, opportunity, and motivation impact their future behavior; and their behavior in turn impacts their capability, opportunity, and motivation.

Similarly, Han and Weiss (2005) highlight the impact of individual-level factors on the sustainment of school-based mental health programs in their process model of sustained implementation, as the decision to implement a practice ultimately falls on each individual teacher, even in the most ideal implementation contexts. Like Michie et al. (2011), Han and Weiss (2005) consider a teacher's motivation to engage in a behavior as an essential component to the sustainability of a program. In their process model, motivation sits alongside skill; both of which are impacted by engaging in the behavior, experiencing success, and attributing the success to the new behavior (independently or with the support of a coach). Furthermore, the Health Action Process Approach (HAPA; Schwarzer, 2008)—another behavior change model that seeks to narrow the gap between behavioral intentions and subsequent behavior—highlights the important role of intentions on future action. In the HAPA, intentions and the factors likely impacting intentions (i.e., self-efficacy, outcome expectancies, perception of risk, barriers, and resources) are situated within the preintentional motivation phase, during which behavioral intentions are generated.

The importance of considering an individual's intentions when attempting to understand and possibly change subsequent behavior is evident, even when based solely on this limited sample of behavior change literature. Keeping this in mind, while recognizing the variety of contextual factors highlighted in and across models thought to impact intentions and behavior, all future explorations of ECSE teachers' data collection practices should be grounded in a theory that focuses on teachers' intentions to engage in data collection while at the same time considering the many contextual barriers unique to data collection in ECSE. In addition to addressing the limitations previously discussed, creation of a scale grounded in such a theory would be instrumental in the prediction and explanation of teachers' level of coherence with recommended (and legally required) practices when it comes to data collection. Additionally, creation of a theory-driven scale rooted in teachers' perspectives of factors known to impact data collection would be valuable to the development and evaluation of implementation supports at the pre- and in-service level.

Theory of Planned Behavior. The Theory of Planned Behavior (TPB; Ajzen, 1985) may help to not only aid in the organization of future research in this area, but also in the creation of a tool used to predict and explain data collection practices by measuring teachers' *behavioral intentions*. Conceptualized by Ajzen as "indications of how hard people are willing to try, of how much of an effort they are planning to exert" (1991, p.181), behavioral intentions are similar to how others have described motivation. Viewed as the key determining factors of ensuing behavior, intentions are ultimately determined by an individual's beliefs (see Figure 2). *Behavioral beliefs* represent one's beliefs about a behavior and their evaluative judgements regarding the outcomes of the behavior, which form a person's *attitudes* toward the behavior. *Normative beliefs* include a person's beliefs about the expected norms associated with a behavior (also referred to as social pressures) and their motivation to behave in accordance with these norms, which is thought to establish one's *subjective norms*. C*ontrol beliefs*, or an individual's "perception of the ease or difficulty of performing the behavior of interest" (Ajzen, 1991, p. 183), establishes their *perceived behavioral control* (commonly referred to as self-efficacy in other conceptualizations). In addition to its indirect impact on behavior through influencing intentions, an individual's perception of their behavioral control (given it is an accurate measure of the factors that actually serve as facilitators or barriers to their behavior) is thought to also directly impact behavior by moderating the predictive power intentions have on behavior.

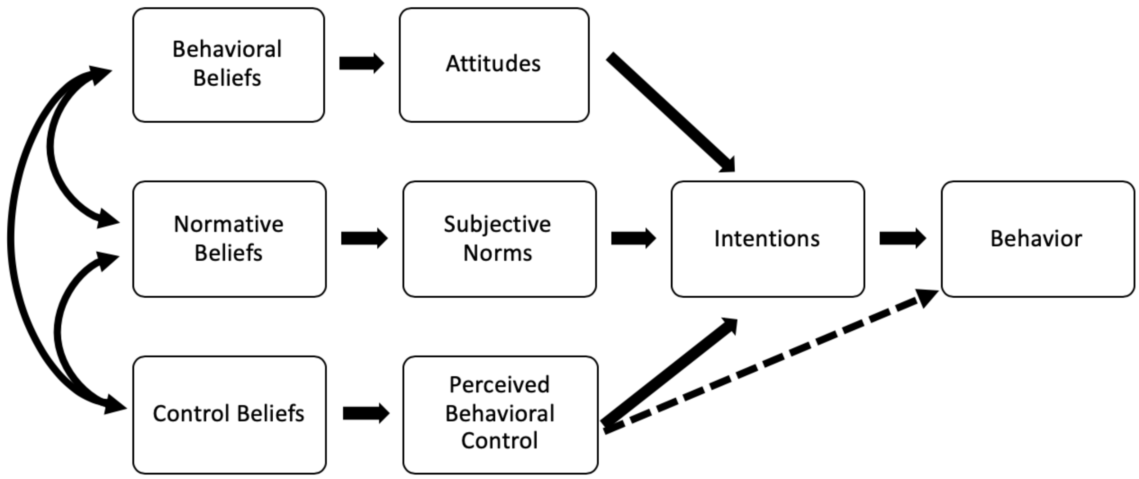

Figure 2. The Theory of Planned Behavior. This figure illustrates the process from beliefs to behavior, highlighting the direct impact of perceived behavioral control on both intentions (solid line) and behavior (dashed line).

Over the past three decades, the TPB has been extensively applied in the health sciences field to better understand and predict the behavior of individuals across various populations (Armitage & Conner, 2001). Results from a meta-analysis of empirical examinations of the TPB—including a wide range of participants, behaviors, and contexts—found (a) belief-based measures of attitude, subjective norm, and perceived behavioral control to be strongly correlated with direct measures of these constructs; (b) attitude, subjective norm, and perceived behavioral control accounted for an average of 39% of the variance in behavioral intentions; and (c) intentions and perceived behavioral control accounted for an average of 27% of the variance in the predicted behavior; thus providing support for the prediction of behavior via the measurement of these beliefs (Armitage & Conner, 2001).

More recently, the TPB has been proven useful to the field of education by identifying critical factors impacting teachers' behaviors in the classroom (MacFarlane &

Woolfson, 2013; Ruble et al., 2018; Yan & Sin, 2014). To highlight the potential organizing utility of this theory, both in uncovering factors influencing teachers' data collection practices and explaining relations among factors, as well as in the creation and evaluation of implementation supports, a discussion of empirical support from the field of early childhood is warranted. In addition to exploring how ECSE teachers' attitudes, subjective norms, perceived behavioral control, and intentions have been measured, the limitations associated with the measurement of these constructs will be discussed.

Measurement of attitudes. According to the TPB (Ajzen, 1985), one's attitudes toward a behavior is formed by their beliefs related to the behavior and the consequent behavioral outcomes (both positive and negative in nature). In measuring special education teachers' attitudes about data collection, Ruble et al. (2018) created a 17-item survey referred to as TIDE (Teacher Intention Toward Data Collection Efforts). Three of the 17 items were specific to teachers' attitudes and focused on the degree to which teachers believe (a) data collection to be important, (b) collecting data will help their students' meet IEP goals, and (c) students who meet their IEP goals are more successful. Response options included a 6-point rating scale from *not important* or not *completely false* to *extremely important* to *completely true*. Results indicated that teachers' attitudes toward data collection correlated positively with teachers' self-reported intentions to collect data ($r = .46$, $p < .001$).

Measurement of subjective norms. In addition to teachers' attitudes, included in the TPB is the notion that a person's normative beliefs—beliefs about the expected norms and their motivation to behave in accordance with these norms—determine their subjective norms (Ajzen, 1985). Because an individual "can hold normative beliefs with respect to more than one referent individual or group" (Ajzen, 2012, p. 5), it is important to explore teachers' perceptions of the beliefs held not only by their administrators, but also by their colleagues and students' parents. With respect to this, Ruble and colleagues (2018) measured teachers' subjective norms using nine survey items focused on teachers' perceptions of their administrators', coworkers', and students' parents' (a) beliefs regarding the importance of data collection, (b) willingness to support their data collection, and (c) attention toward collected data. Scored using a 6-point rating scale from *not important*, *not at all likely*, or *strongly disagree,* to *extremely important, extremely likely, or strongly agree*, responses to these items represented teachers' subjective norms. Results indicated that teachers' normative beliefs also correlated positively with teachers' self-reported intentions to collect data ($r = .59$, $p < .001$).

Measurement of perceived behavioral control. While a teacher's desire or intent to perform certain behaviors might be predictive of subsequent behavioral performance (e.g., a teacher who intends to read a story during group time is very likely to read that story during group time), actual performance that is not aligned with a teacher's intentions is not uncommon. Whether the behavioral intention represented a difficult or unrealistic behavior, or an unexpected event impacted the teacher's ability to perform the behavior, it is clear to see that a variety of factors—both internal and external—can break the chain from intending to behave in a certain way and actually demonstrating the behavior. According to the TPB, a person's perception of the ease or difficulty associated with a behavior (i.e., control beliefs) determines their perceived behavioral control, and is often driven by past experiences and perceived barriers (Ajzen, 1991). Therefore, when controlling for behavioral and normative beliefs, a teacher who views the frequent collection of IEP data as a simple and straightforward task is more likely to implement this type of data collection than a teacher who views it as a challenging task.

Ruble et al. (2018) measured teachers' perceived behavioral control using four survey items, all consisting of the same 6-point rating scale from *strongly disagree* to *strongly agree*. Items measuring behavioral control focused on the following hypothesized barriers to data collection in ECSE: (a) too little time, (b) too much to do, (c) unclear measurement systems, and (d) too many students. Results indicated that teachers' control beliefs were correlated positively with teachers' self-reported intentions to collect data, but the relationship between the two variables was relatively weak ($r = .32$, $p < .05$).

In addition to perceived behavioral control, Ruble et al. (2018) measured teachers' self-efficacy related to data collection by asking teachers to rate themselves on a 100-point scale (0 = cannot do at all; 100 = certain can do) based on their ability to collect data for the purposes of progress monitoring as well as their ability to use the data to re-evaluate IEP goals and objectives. These two items were taken from the Autism Self-Efficacy Scale for Teachers (ASSET; Ruble, Toland, Birdwhistell, McGrew, & Usher, 2013) and were administered in addition to the 17-item TIDE. Results indicated that self-efficacy was correlated with teachers' intentions to a greater extent than perceived behavioral control ($r = .55$, $p < .001$), though the correlation with their actual data collection practices was not significant.

Measurement of behavioral intentions. As previously described, an individual's intention (or motivation) to engage in a behavior is ultimately determined by their attitudes, subjective norms, and perceived behavioral control. In turn, intentions are thought to predict and/or explain future behavior. Ruble et al. (2018) measured teacher's intentions to engage in future IEP data collection using one survey item: "I intend to keep data over the next two weeks." The response options consisted of a six-point scale ranging from not at all likely to extremely likely. Results indicated that individuals' responses to this single item did not produce a significant correlation with their actual data collection practices ($r = .25$, $p = .10$).

Measurement limitations. Though Ruble and colleagues' (2018) exploration of teachers' intentions to engage in IEP data collection represents movement toward a more organized examination of data collection behaviors—one based on a well-defined and researched theory—there are several limitations related to the creation of the questionnaires utilized as well as the methods employed to analyze the data. First, the items included in the TIDE may not have been sufficient in defining each construct. For example, a potential explanation for the weak correlation between perceived behavioral control and intentions may be that administrative support (e.g., access to training, adequate classroom staffing, time for planning) was measured and analyzed as a separate construct (which was also found to correlate with teachers' intentions; $r = .35$, $p < .05$), even though these variables impact the ease or difficulty associated with data collection. Furthermore, keeping in mind that data collection is an intricate process, teachers' perception of their capacity related to each component of data collection (writing clear and measurable IEP objectives, creating/modifying measures, assigning data collection responsibilities across each student's educational teams, etc.) should be examined. A final example of insufficient coverage was the measurement of behavioral intentions using one single item.

Next, while the internal consistency of the TIDE given the sample used was reported, no additional methods providing construct-related validity evidence (e.g., think alouds, factor analysis, item response theory modeling, etc.) were described. That is, it was not convincing that the items created were accurate or complete measures of the constructs targeted. Finally, ordinal data were collected, yet item and scale means were not only reported, but were treated as continuous variables in subsequent correlation and standard linear regression analyses. Because there is no way to determine whether the intervals between each rating (e.g., strongly disagree, disagree, agree, strongly agree) on a scale are equal—that is, we cannot assume the difference between ratings of disagree and agree is the same as the difference between ratings of agree and strongly agree—it is inappropriate to treat ordinal data derived from rating scales as anything other than ordinal data without the use of subsequent analyses. Additionally, calculating the means of individuals' responses across a set of items can be very misleading, as it does not account for varied responses to items designed to tap different aspects of the construct. Item response theory (IRT) modeling provides an approach to overcoming both of these barriers commonly faced when working with survey data.

Conclusion

Regardless of the legal requirements and empirical evidence, a discrepancy exists between what is known and what is practiced in relation to data collection (Brawley & Stormont, 2013; Cook et al., 1991; Fuchs et al., 1982; Sandall et al., 2004; Wesson et al., 1984). Compared to the exploration of progress monitoring in K-12 settings, data collection in ECSE has been relatively ignored, which is concerning given the unique instructional contexts in which ECSE teachers are expected to collect data. Though negligible in size, the literature base suggests that ECSE professionals generally agree that data collection is important for a variety of reasons. A discrepancy exists, however, between their beliefs and their actual practices (Brawley & Stormont, 2013; Ruble et al., 2018; Sandall et al., 2004). Given the importance of data collection, a number of supports have been developed to facilitate data collection in ECSE, from structured frameworks for collecting and analyzing data to numerous data collection tools and various electronic technologies. Nevertheless, the sporadic, unstructured, and relatively subjective data collection methods continue to persist (Brawley & Stormont, 2013; Sandall et al., 2004).

In an attempt to explain the gap between research and practice when it comes to the frequent collection of IEP data in ECSE, a variety of internal and external factors potentially serving as barriers to data collection have been explored. In addition to a teacher's knowledge and skills in relation to data collection, internal factors impacting one's motivation to engage in data collection include a teacher's perceptions of (a) their capacity, (b) the importance of the IEP and subsequent data collection, (c) the practicality of ongoing data collection, and (d) the demand to modify their existing data collection practices. External factors hypothesized to influence teachers' data collection practices include leadership support, collegial support, and the instructional context in which data collection takes place.

Though numerous barriers to data collection have been identified that help to illuminate possible reasons for the gap between research, teachers' reported beliefs, and teachers' reported practices, the relations between these factors are unknown due to the organization and clarity of the literature. In addition to the absence of a theory-driven framework guiding the research in this area, inconsistently and only partially defined constructs (i.e., beliefs and data collection) are not uncommon. Furthermore, commonly employed data analysis procedures have raised questions about the validity and reliability of the conclusions drawn. As a result, it is not clear whether a discrepancy truly exists between what teachers believe to be important and what they reportedly practice, or whether this gap has been influenced by unsatisfactory measurement.

To assist in the creation of implementation supports that increase ECSE teachers' ability to not only benefit from existing facilitators of data collection, but also to commit to overcoming common barriers to data collection, it is imperative that these limitations are addressed in future research. The TPB (Ajzen, 1985) has the potential to support both our understanding and prediction of data collection practices in ECSE by focusing on a critical individual-level factor linked to implementation—teachers' behavioral intentions. Additionally, greater attention toward improving the measurement procedures included in previous applications of this theory will aid in the development of a tool that can be used in future explorations of teachers' intentions to collect IEP data; one that produces data that could be more easily and accurately summarized and interpreted within and across studies.

Chapter Three – Methods

Previous applications of the Theory of Planned Behavior (TPB; Ajzen 1985) have demonstrated its promising capability in providing the organization and clarity that is imperative to future explorations of teachers' IEP data collection practices in ECSE. Recognized as a framework that encompasses a set of fundamental beliefs known to impact individuals' intentions (similarly referred to as commitment or motivation) to engage in a behavior, the TPB also recognizes individuals' perceptions of a variety of additional factors that have been found to impact implementation. Given these qualities and considering the limitations previously discussed, a systematic replication of Ruble et al. (2018) is warranted, with an enhanced approach to the measurement of each construct.

In this application of a cross-sectional survey design, particular attention will be given to the inclusion of methods known to increase the validity of the interpretation and use of data generated from teachers' completion of the IEP Data Collection Intentions Scale (IDCIS). Following a description of the study's specific aims and procedures employed to recruit participants, the IDCIS's interpretation and use argument (IUA) will be presented, which was used to guide all subsequent procedures described in this chapter including the development and refinement of the IDCIS and data analysis. With a focus on improving Ruble and colleagues' measurement of attitudes, subjective norms, perceived behavioral control, and intentions, item response modeling was employed to allow for the precise location of individuals on each trait continuum in future applications of the IDCIS.

Study Purpose

The purpose of this study was to evaluate the dimensionality and quality of the IEP Data Collection Intentions Scale (IDCIS)—a scale grounded in the TPB—based on its intended interpretations and uses. In doing so, answers to the following three research questions were explored:

1. To what extent, if at all, does the IDCIS represent four distinct constructs, including teachers' attitudes, subjective norms, perceived behavioral control, and behavioral intentions toward IEP data collection?
2. Which items serve as quality indicators of each construct, such that valid and reliable inferences about teachers' IEP data collection intentions—regardless of level—can be made using the resulting data?
3. Does the Theory of Planned Behavior serve as an appropriate theoretical model to measuring teachers' IEP data collection intentions, such that teachers' attitudes, subjective norms, and perceived behavioral control explain a significant amount of the variance in ECSE teachers' intent to engage in future IEP data collection?

Interpretation and Use Argument for the IDCIS

All study procedures—from initial scale development to data analysis—were grounded in the scale's interpretation and use argument (IUA). As such, a brief explanation of the IUA is warranted prior to the description of subsequent methods. Regarded as the first step in the argument-based approach to validation, the IUA consists of a set of claims about the proposed score interpretation and uses (Kane, 2013). In addition to these claims, the IDCIS's IUA also includes claims regarding the constructs being measured. Claims regarding the constructs of attitudes, subjective norms, perceived behavioral control, and intentions included the following: 1) supported by the TPB (Ajzen, 1985), teachers' intentions to engage in future IEP data collection are determined by their attitudes, subjective norms, and perceived behavioral control, which are all correlated; and 2) given the organization of constructs within the TPB, this scale is ultimately a measure of teachers' IEP data collection intentions, thus providing information regarding how much of an effort they are planning to exert when it comes to collecting IEP data. Claims regarding conclusions about IDCIS score interpretations and uses are that the IDCIS produces scores that are precise measures of teachers' attitudes, subjective norms, and perceived behavioral control related to IEP data collection and teachers' intentions to engage in future IEP data collection, such that meaningful information about these constructs can be used in the development and/or modification of implementation supports.

In validating the proposed interpretation and uses of the IDCIS (the second step in the argument-based approach), it is imperative to consider the sequence of inferences required to move from an individual's responses to the items on the IDCIS to the conclusions drawn about the individual. According to Kane (2013), IUAs for indicators of theoretical constructs typically include at least three inferences: 1) scoring, 2) generalization, and 3) extrapolation. In the context of the IDCIS, the scoring inference includes assumptions about the creation and application of scoring procedures, while the generalization inference assumes an individual's observed score represents their average performance, should they have completed the scale on a number of occasions. The extrapolation inference includes, what Kane describes as a "more-ambitious leap" and assumes the generalized indicators represent an individual's actual ability in regard to each latent trait being measured (2013, p.28). This sequence of inferences, as seen in Figure 3, served as the foundation for all subsequent validation procedures. As such, specific forms of validity evidence used to support each inference will be highlighted in conjunction with the descriptions of all procedures moving forward.

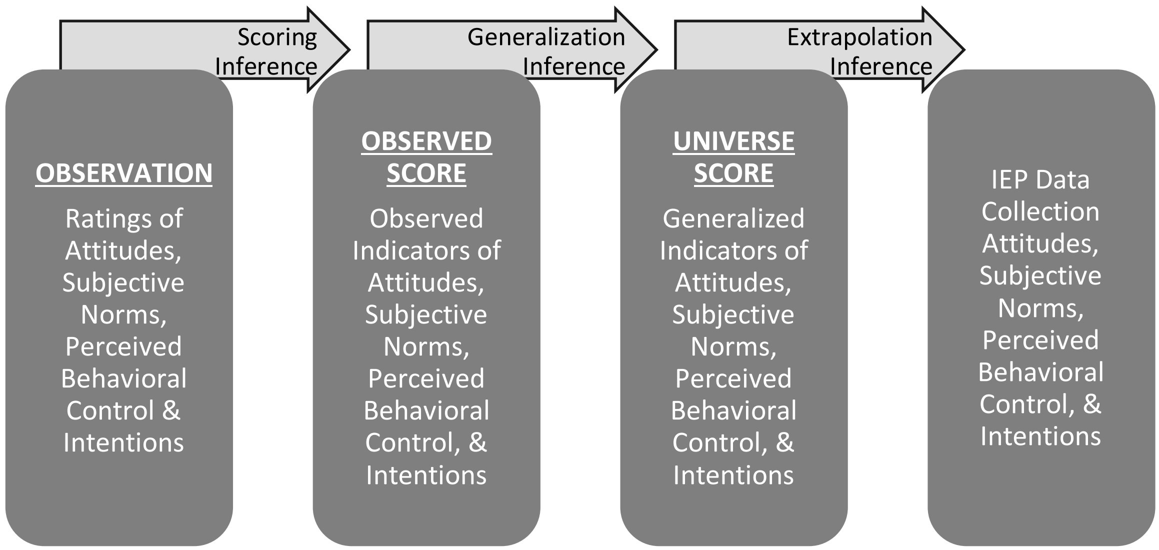

Figure 3. Sequence of Inferences Guiding Validation Procedures

Initial Scale Development

Development of the IDCIS was firmly grounded in construct modeling, an instrument development approach rooted in item response modeling (Wilson, 2005). Keeping in mind the most proximal use of this scale, which is to provide precise measures of teachers' IEP data collection beliefs and intentions, it was important to keep item response modeling top of mind starting from ground zero—scale development. Just as a ruler can be used to accurately measure an object's length (thus allowing for a group of objects to be meaningfully compared based on their respective lengths), item response modeling has the ability to support the creation of a tool that produces accurate measures of latent variables that are inherently hard to measure, such as attitudes, subjective norms, perceived behavior control, and intentions. As described by Wilson (2005), construct modeling is comprised of four building blocks including 1) the development of a construct map, 2) item design, 3) specifying the outcome space, and 4) applying the measurement model. Following is a detailed description of scale development procedures associated with each of these building blocks.

Step One—Development of construct maps. Defining the construct to be measured and mapping out the full range (e.g., low to high, weak to strong) of the construct is the first step of scale development grounded in item response modeling (Wilson, 2005). The IDCIS was ultimately designed to measure teachers' *intentions* to engage in future IEP data collection. Guided by the Theory of Planned Behavior (TPB; Ajzen, 1985), however, teachers' (a) *attitudes*, (b) *subjective norms,* and (c) *perceived behavioral control* will also be measured, as it is these constructs that have been found to explain variances in intentions. A review of existing literature provided initial support to the defining of constructs, followed by an expert review.

As recommended by Francis and colleagues (2004), *attitudes* will be conceptualized as teachers' perceptions regarding the utility of IEP data collection (similar to what has been formerly described as "importance") as well as their overall feelings associated with the act of engaging in IEP data collection. Instead of asking a few general questions about the importance of IEP data collections, several items that tap teachers' perceptions of data collection's capacity to improve various aspects of their jobs will be included in hopes to produce a more precise measurement of this construct. Consistent with the TPB literature, the construct of *subjective norms* will include teachers' perceptions of how important people (e.g., leadership, coworkers, parents) view IEP data collection. More common than not, this construct has historically been measured using a single item and has been found to have a weak association with behavioral intentions (Armitage & Conner, 2001). Therefore, the conceptualization of subjective norms will be expanded in the current study to include teachers' perceptions of the observed behaviors of important people (i.e., what people talk about, attend to, and reward), as we know these behaviors impact organizational climates, and thus individuals' behaviors (Schein, 2004).

As recommended by Francis et al (2004), *perceived behavioral control* will be measured by two sets of items; one tapping teachers' self-efficacy related to IEP data collection behaviors and one related to the controllability of their IEP data collection behaviors (i.e., what barriers they face and how much power they have in overcoming each barrier). Improving upon previous measures of control beliefs, this scale will include self-efficacy items spanning a variety of data collection behaviors (e.g., writing clear and measurable IEP objectives, modifying existing data collection measures, creating new data collection forms) and will include an exhaustive list of barriers.

Finally, the construct of *intentions* will be conceptualized similar to how Armitage and Conner (2001) theorized self-predictions, which was the extent to which participants intended to engage in a future behavior. The recommendation of Eccles et al. (2004) to include three items—one representing expectations ("I expect to…"), another representing desires ("I want to…"), and the final representing intentions ("I intend to…")—was not followed, as responses to these items have been extremely consistent (Armitage & Conner, 2001). Additionally, to boost coverage by limiting the time commitment involved in survey completion, the use of vignettes in measuring intentions (e.g., MacFarlane & Woolfson, 2013) was decided against.

After defining each construct, construct maps were created to ensure items were designed to tap each level of the construct. For example, when defining *attitudes*, the construct was conceived with the full range of attitudes in mind; from an extremely negative attitude toward IEP data collection (e.g., someone who believes data collection is not at all useful, data collection does not improve any aspect of their job, and the thought of data collection produces negative feelings) to an exceptionally positive attitude toward IEP data collection (e.g., someone who believes data collection is imperative, data collection improves all aspects of their job, and the thought of data collection produces positive feelings). The empirical support for the use of the TPB (Ajzen, 1985) in measuring intentions as well as the application of construct maps in scale development served as validity evidence based on scale content—both construct representation and coverage—that supported the extrapolation inference.

Step Two—Item design. Preceded by the defining and mapping of each construct, item design involves the process of writing the items that will serve as indicators of each construct (Wilson, 2005). In addition to ensuring that each item is aligned with and represents a different level of each construct, survey quality—including the wording of individual items, their order within the survey, and the survey's visual design—was the main focus of the item design phase in order to minimize measurement error due to inaccurate responses. As well as attending to how items have been written in previous studies, evidence-based item-writing guidelines recommended by Dillman et al. (2014) were heavily relied on during this phase of scale development. Specific guidelines endorsed by Dillman and colleagues (2014) that were followed during item design include, but are not limited to: (a) the utilization of branching within Qualtrics to ensure all participants are prompted to respond to only items that are individually applicable to them; (b) keeping items as concise as possible; (c) explicitly defining important construct-specific terminology; (d) including both ends of unipolar and bipolar scales in the item stem, so as not to give priority to one over the other; (j) organizing items in a meaningful way; and (k) beginning with an item thought to be salient to most survey respondents. Finally, to decrease the cost associated with participation, particular attention was paid toward reducing the complexity of the survey by limiting the survey length, creating a survey layout that is easily navigated, and limiting the amount of personal information collected. Solid item construction rooted in evidence-based item writing guidelines served as one form of validity evidence supporting the scoring inference.

Step Three—Outcome space. While attending to how items were constructed, it was imperative to be mindful of the outcome space, or the categorization of response options and their subsequent scoring, as this will serve as the basis of inferences made about each individual's level of measured constructs (Wilson, 2005). Similar to item design, planning of the outcome space was supported by a thorough review of the literature and evidence-based item writing guidelines. In addition to ensuring that an exhaustive list of responses was generated and that response options were appropriately scaled to cover various levels of each construct as suggested by Wilson (2005), the following guidelines recommended by Dillman et al. (2014) were followed: (a) ensuring response options were mutually exclusive; (b) including forced-choice items in lieu of check-all-that-apply; (c) avoiding an odd number of response options that include a neutral response; (d) labeling all response options along a scale; and (e) balancing the order of response options to deter primacy or recency effects. Attention toward these guidelines when designing the outcome space created additional validity evidence supporting the scoring inference.

Step Four—Measurement model. The measurement model—the final building block in construct modeling—is what allows us to identify an individual's location on the construct map based on their scored responses to each item (Wilson, 2005), thus supporting the scoring inference. Given its ability to produce precise and equal-interval measures of constructs, item response theory (IRT) modeling will be used to ensure that future IDCIS scores are dependent on each individual's level of the construct, rather than on the sample obtained or the specific items included on the scale. Specific procedures associated with item response modeling will be described within the data analysis section below.

Scale Refinement

Following initial scale development, the IDCIS included seven items related to attitudes, seven items related to subjective norms, 26 items related to perceived behavioral control (10 aligned with self-efficacy, and the remaining 16 aligned with controllability), and 4 items related to intentions. With the goal of creating more validity evidence necessary to support the IDCIS's IUA, the scale underwent an iterative refinement process, which included multiple rounds of feedback and subsequent modification. Refinement procedures included expert reviews, a think aloud, and a small pilot study.

Expert reviews. Three rounds of expert reviews were conducted to ensure appropriate scale content (i.e., representation of each construct and content coverage) as well as to ensure adherence to evidence-based item-writing guidelines, thus producing additional validity evidence supporting the extrapolation inference. All review correspondences were via email; experts were sent a brief background on the creation of the IDCIS as well as a copy of the initial scale and were asked to provide feedback in the form of electronic comments and/or track changes. The first round of reviews included two ECSE content experts, both of whom were doctoral students in the Special Education Program at a local university and had practical and research experience in ECSE. The ECSE content experts were asked to focus their attention on content coverage, with a specific eye on whether the IDCIS items spanned the full range of each construct. Specifically, they were asked to ensure the scale included items that would be easily endorsed by teachers who exhibit very low to extremely high levels of each construct.

The second round of reviews included two TPB (Ajzen, 1985) content experts from a local university—one doctoral student in the School Psychology Program and one faculty member in the Special Education Program. Again, both experts had experience in practice and research; one of which who also had an extensive background in ECSE. The TPB content experts were asked to focus on whether the TPB constructs were appropriately represented across each set of items. The final round of reviews included a faculty member in the Quantitative Methods in Education Program at a local university whom served as the technical expert. This individual was asked to pay particular attention to the degree to which evidence-based item-writing guidelines were followed.

Think aloud. Following scale refinement based on feedback generated by the expert reviews, a think aloud was conducted in order to better understand individuals' thought processes as they read and responded to each survey item, creating additional validity evidence supporting the scoring inference. The think aloud—also known as a cognitive interview (Willis, 2004)—took approximately 90 minutes to complete and was conducted after all expert reviews, yet prior to piloting procedures. The think aloud participant was a licensed ECSE teacher who had approximately 25 years of experience teaching young children with delays/disabilities in classroom settings, but who did not meet all criterion for study participation due to their current position as a district-wide instructional coach.

As recommended by Dillman and colleagues (2014), the think aloud began with a description of what the participant should expect followed by the practice question, "How many windows do you have in your house?" As the participant began thinking about the number of windows in their house, they were frequently prompted to think out loud and explain their reasoning, which further highlighted the expectations. When the process was clear, the participant was given a paper copy of the survey and the think-aloud began. The participant was prompted as needed to verbalize all necessary information and their thoughts related to each item were recorded on paper for the duration of the interview. Following the think-aloud, the IDCIS was modified and shared with the think-aloud participant for review to ensure the modifications were appropriately reflected their thoughts during the think aloud.

Pilot Study. To gain additional information about individuals' response processes following the think-aloud and to provide further validity evidence supporting the scoring inference, a small pilot study was conducted using the most up-to-date version of the IDCIS. With the goal of maximizing the number of possible study participants, only ECSE teachers who did not meet all inclusion criterion were recruited for participation in the pilot study. Recruitment procedures consisted of a single email distributed to personal acquaintances of the researcher explaining the need for individuals to test the survey and provide feedback on the survey. Of the 30 individuals who were contacted, 25 were licensed and currently practicing ECSE teachers who were not presently working with children on IEPs in classroom settings, and five were graduate students in an ECSE licensure program who were currently working with young children on IEPs in classroom settings.

In total, 17 individuals completed the survey, resulting in a response rate of 57%. Because the pilot participants were not eligible for study participation given their current role as a licensed ECSE teacher or their licensing status, they were arguably different from the population the IDCIS was developed for. Therefore, the purpose of the pilot round was to ensure (a) items were clearly written and easy to understand; (b) the survey was easily accessible and properly functioning across all preferred electronic devices; and (c) items were coded properly and (d) data produced by Qualtrics were formatted to facilitate the intended analyses.

Final Instrumentation

At the conclusion of all refinement procedures, the final version of the IDCIS used in the present study included 53 items; eight items serving as indicators of attitudes, eight items serving as indicators of subjective norms, 30 items serving as indicators of perceived behavioral control (10 aligned with self-efficacy, and the remaining 20 aligned with controllability), and seven items serving as indicators of intentions. Construct maps highlighting each included item as well as the hypothesized placement of items along each trait continuum are displayed in Figure 5 through Figure 8. All items included 4 response options, which were coded such that a score of 1 represents the lowest level of the construct and a score of 4 represents the highest. Some items were displayed as individual questions with a corresponding set of response options, while other items were grouped into matrices and included a shared set of response options. After an additional 11 demographic questions were included to gain important information about respondents, the survey completed by study participants took approximately 20 minutes to complete. See Appendix A for a complete version of the survey (including all response options) administered in this study.

HIGH Level of Attitudes

A5. To what extent, if at all, do you agree that IEP DC improves the quality of your instruction?
A8. To what extent, if at all, do you agree that IEP DC improves the quality of your students' outcomes?
A3. To what extent, if at all, do you agree that IEP DC improves the quality of your IEP objectives?
A7. To what extent, if at all, do you agree that IEP DC improves the quality of your communication with parents?
A6. To what extent, if at all, do you agree that IEP DC improves the quality of your accountability to others?
A4. To what extent, if at all, do you agree that IEP DC improves the quality of your progress reporting?
A2. How useful, if at all, is IEP DC?
A1. How important, if at all, is IEP DC?

LOW Level of Attitudes

Figure 5. Construct Map for Attitudes.

HIGH Level of Subjective Norms

SN8. How often, if ever, do you observe someone in an ECSE leadership role in your district acknowledge you for your IEP data collection efforts?
SN6. How often, if ever, do you observe someone in an ECSE leadership role in your district look at your IEP data.
SN5. How often, if ever, do you observe someone in an ECSE leadership role in your district communicate with you about IEP DC?
SN9. How often, if ever, do you observe someone in an ECSE leadership in your district communicate with you about IEP DC?
SN4. How often, if ever, do you observe your coworkers engaging in IEP DC?
SN7. How often, if ever, do you observe someone in an ECSE leadership role collect data highlighting student, staff, or program performance?
SN3. How important, if at all, is IEP DC to your students' parents?
SN1. How important, if at all, is IEP DC to ECSE leadership in your district?
SN2. How important, if at all, is IEP DC to your coworkers?

LOW Level of Subjective Norms

Figure 6. Construct Map for Subjective Norms.

HIGH Level of Self-Efficacy

SEC5. How confident, if at all, are you in your ability to consistently carry out your plan, such that IEP data are collected on a daily basis?
SED5. How difficult, if at all, is it to consistently carry out your plan, such that IEP data are collected on a daily basis?
SEC4. How confident, if at all, are you in your ability to design a plan for the daily collection of IEP data?
SED4. How difficult, if at all, is it to design a plan for the daily collection of IEP data?
SEC3. How confident, if at all, are you in your ability to create new assessment tools that meet your IEP DC needs?
SED3. How difficult, if at all, is it to create new assessment tools that meet your IEP DC needs?
SEC2. How confident, if at all, are you in your ability to modify existing assessment tools to meet your IEP DC needs?
SED2. How difficult, if at all, is it to modify existing assessment tools to meet your IEP DC needs?
SEC1. How confident, if at all, are you in your ability to write clear and measurable IEP objectives across all developmental domains?
SED1. How difficult, if at all, is it to write clear and measurable IEP objectives across all developmental domains?

LOW Level of Self-Efficacy

HIGH Level of Controllability

How often, if ever, do these factors decrease your ability to engage in daily IEP DC?
B1. Availability of time to plan for IEP DC
B2. Availability of time to collect IEP data
B9. Adult-to-child ratio in the classroom
B14. Number of students on your caseload
B17. Other DC requirements aligned with program, school, district, or state initiatives
B15. Your role in the planning and delivery of instruction and/or support within the classroom
B8. Nature of classroom routines/activities (i.e., balance of highly structured and adult-led one-on-one activities vs. minimally structured and child-led group activities)
B13. Range of developmental levels across all students on your caseload
B4. Access to commercially-available paper-pencil tools that meet your IEP DC needs
B5. Access to commercially-available electronic tools that meet your IEP DC needs
B10. Student's developmental levels in the classroom
B7. Classroom curriculum
B6. Classroom type (i.e., self-contained vs. inclusive)
B12. Number of IEP objectives on your students' IEPs
B11. Clarity of IEP objectives on your students' IEPs
B16. Availability of professional development geared toward IEP DC in ECSE
PBC1. To what extent, if at all, do you agree that the quality of your students' IEP objectives is entirely up to you?
PBC2. To what extent, if at all, do you agree that the choice of tools you use to collect IEP data is entirely up to you?
PBC3. To what extent, if at all, do you agree that whether or not you engage in IEP DC is entirely up to you?

LOW Level of Controllability

Figure 7. Construct Map for Perceive Behavioral Control.

HIGH Level of Intentions

I7. Over the next month, to what extent, if at all, do you intend to observe students while formally documenting the frequency and/or duration of their skills/behaviors AND the type and/or frequency of adult prompts provided?

I6. Over the next month, to what extent, if at all, do you intend to observe students while formally documenting the frequency and/or duration of their skills/behaviors?

I2. To what extent, if at all, is IEP DC a high priority for you over the next month?

I1. To what extent, if at all, do you plan to collect IEP data across all students on your caseload over the next month?

I5. Over the next month, to what extent, if at all, do you intend to observe students while taking informal notes about their skills/behaviors?

I4. Over the next month, to what extent, if at all, do you intend to talk with others (i.e., educational assistants, related service providers, and parents) about their observations of students' skills/behaviors while taking informal notes?

I3. Over the next month, to what extent, if at all, do you intend to reflect on your memory of students' skills/behaviors while writing down informal notes?

LOW Level of Intentions

Figure 8. Construct Map for Intentions.

Recruitment

Following approval from the University Institutional Review Board, all individuals who met the following inclusion criteria were recruited to participate: those who (a) held a current ECSE license, (b) were currently working as an ECSE teacher, and (c) had been working with at least one student with an Individualized Education Program (IEP) in a classroom setting for the past two months. A fairly large sample size was needed in order to complete the necessary analyses; therefore, the coverage goal of this study was to census the entire population of ECSE teachers in the state of Minnesota. Because a list of all individuals in the target population does not exist, two databases were used to estimate the population size. Given there were approximately 2000 ECSE teachers employed across the state during the 2017-2018 school year (MN PELSB, 2018) and approximately 75% of students in ECSE were receiving Part B (special education for individuals ages 3 through 21) services (MDE, 2018), it was estimated that approximately 1500 (2000 x .75 = 1500) ECSE teachers across the state had the *potential* to meet all inclusionary criteria; though a portion of these teachers likely serve on evaluation teams or as professional development coaches and thus would not meet the inclusion criteria.

ECSE leaders. Given there was no direct access to the target population, district-level ECSE leadership (coordinators, supervisors, etc.) served as the intermediator to the population. Compilation of a complete list of ECSE leadership representing all organizations (i.e., independent districts and schools, special districts and schools, special education cooperatives, intermediate districts and schools, service cooperatives, charter schools, and education districts) employing ECSE teachers across the state was supported by the following data sources: (a) the "2018-2019 ECSE Leaders" list provided by the Minnesota Department of Education (MDE), which includes the names, tiles, and email addresses of all ECSE leaders across 351 educational organizations; and (b) the "2017-2018 Licensed Headcount FTE by Assignment Code" list downloaded from the Minnesota Professional Educator and Licensing Standards Board website (MN PELSB, 2018), which includes the name of each district across the state as well as the number of licensed ECSE teachers employed by each district. The final list included 139 ECSE leaders who represented 336 organizations that employed at least one ECSE teacher during the 2017-2018 school year.

Upon compilation of this list, an email was sent to all 139 leaders to assess their willingness to participate in the study. Participation of ECSE leaders consisted of forwarding three emails—an initial email inviting participation and two reminder emails—over a period of approximately three weeks to all individuals in their organization who met the inclusion criteria. To minimize nonresponse error, a common source of error present in most survey research, the initial call for participation of ECSE leaders included a well-defined purpose, clearly linking the leaders' participation to common needs in the field of ECSE in order to increase their motivation to participate. Additionally, participating leaders were promised a summary of the survey results as well as access to an online training on IEP data collection individualized to the barriers faced by ECSE teachers across the state. After eight days, a follow-up email was sent to all non-responders in an attempt to gain access to additional participants. In the end, 42 leaders agreed to participate. The remaining 97 leaders failed to respond.

ECSE teachers. When the list of organizations overseen by the 97 non-participating leaders was reviewed, it was apparent that many of the largest school districts across the state were not represented. Therefore, a web-based search was conducted to obtain the email addresses of ECSE teachers working in the state's 25 largest school districts. ECSE teacher email addresses were located for 22 of these districts, resulting in a list of email addresses for 377 ECSE teachers. Similar to the procedures requested of ECSE leaders, a set of three emails were sent to all teachers on this list over a one-month period—an initial email inviting them to participate and two reminder emails. All emails were sent in the early morning, but were delivered on different days of the week, each approximately one week apart. After sending out the initial call for ECSE teachers' participation, each district's calendar was considered, so as to not send out reminder emails while teachers were out of the office due to spring break.

Again, to limit nonresponse error, the survey itself and all related emails included the researcher's name and contact information (i.e., email address) allowing participants an opportunity to assess the authenticity of the survey and to ask survey-related questions. Additionally, the consent form emphasized the confidentiality and protection of data and all communication with possible survey respondents was designed with professionalism in mind. To increase rewards, all emails included a clear purpose, specifying how the survey would be used and asking for the respondents' help in achieving the researcher's goal. Finally, all participants were given the opportunity to be added into a drawing for one of ten $25 Visa gift cards upon survey completion.

Additional recruitment procedures. In addition to the use of ECSE leaders as intermediators to the population of interest and directly contacting ECSE teachers via their publicly-available email addresses, the survey was advertised at the Minnesota Division for Early Childhood's (DEC) Annual Spring Practitioner Conference via a flyer distributed at a University of Minnesota table and a short verbal advertisement prior to the keynote speaker's presentation. Participants who completed the survey at the DEC conference were given a reusable "Got Data?" tote bag as further incentive to survey completion. Additional participants were recruited through advertisements on social media and personal connections with Minnesota's ECSE community.

Participants

Following all recruitment procedures, the final version of the IEP Data Collection Intention Scale (IDCIS) was distributed via an electronic survey using the online software, Qualtrics (2019). A total of 368 participants responded to a portion or all of the survey. Because the focus of this study was on the creation and validation of the IDCIS, knowing the number of possible participants (i.e., those who received a link to the survey) was not essential, as response rates are only important when making inferences about the population from which the sample was drawn. Though there were 368 total responses, only 333 participants completed the survey in full by answering all questions serving as indicators of attitudes, subjective norms, perceived behavioral control, and intentions.

General demographics. A total of 332 participants responded to the 11 demographic items at the end of the survey. Participants represented all 11 professional development regions, with just over half (51.8%, n = 172) of the participants employed by organizations within Region 11, which is the largest region in the state covering Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington counties. The region with the next highest representation was Region 7, represented by 10.8% (n = 36) of the sample, followed by Region 10 representing 7.5% (n = 25) of the sample. Region 2 represented the smallest number of participants (0.9%, n = 3).

In regard to experience, 28.4% (n = 94) of participants had reportedly been working as *licensed* ECSE teachers for less than four years, 41.1% (n = 137) had been working for 4-13 years, and 17.9% (n = 59) had been working for 14-23 years. Only 12.4% (n = 41) of participants had been working as licensed ECSE teachers for more than 23 years. Most participants were contracted to work 31 to 40 hours each week (94.6%, n = 314) and most were in classroom settings with students on IEPs at least 4 days each week (85.2%, n = 283). Participants most commonly reported caseload sizes of 16-20 students from birth through age six (43.4%, n = 144), followed by caseloads ranging from 10-15 students (28.3%, n = 94), 21-25 students (12%, n = 40), and 6-10 students (9.6%, n = 32). Only 4.2% (n = 14) of the sample reported caseloads with more than 25 students and even fewer participants reported caseloads with less than 6 students (2.4%, n = 8).

Data collection behavior. Included in the demographic items were a series of questions related to participants' IEP data collection behaviors over the past month.

When asked how often they reflected on their “memory of students' skills/behaviors while writing down informal notes,” 59.5% (*n* =197) reported to engage in this behavior *frequently,* while an additional 32.9% (*n* = 109) reported to engage in this data collection behavior *sometimes*. Just over two thirds of the sample (67.7%, *n* = 224) reported to *frequently* document informal notes based on others’ (i.e., educational assistants, related service providers, and parents) reports of students' skills/behaviors, while 30.2% reported to engage in this same data collection behavior *sometimes.* In regard to “observing students while taking informal notes about their skills/behaviors,” 57.7% (*n* = 191) of participants reported to *frequently* engaged in this behavior, while 36.6% (*n* = 121) of participants reportedly engaged in this behavior *sometimes* over the past month.

Though the vast majority of participants reported to document informal notes highlighting student progress, a relatively smaller percentage of the sample engaged in more formal IEP data collection methods. Approximately one fifth of the sample (20.3%, *n* = 67) reported to *rarely* or *never* “[observe] students while formally documenting the frequency and/or duration of their skills/behaviors”, while only 37.8% (*n* = 125) reported to *frequently* engage in this data collection behavior. Finally, when asked about “observing students while formally documenting the frequency and/or duration of their skills/behaviors AND the type and/or frequency of adult prompts provided,” over one quarter of the sample (27.1%, *n* = 90) reported to *rarely* or *never* collect IEP data in this way, while less than one third of participants (30.7%, *n* = 101) reported to *frequently* engage in this type of data collection.

Data Analysis

As with scale development, data analysis was an iterative process involving a sequence of procedures that was repeated in association with each of the four measured constructs. Following an analysis of missing data, each data analysis sequence consisted of four phases including confirmatory factor analysis (CFA; the first stage in structural equation modeling), item analysis, item response theory (IRT) modeling, and structural modeling (the final stage in structural equation modeling). A description of the approaches taken for each set of analyses is provided below.

Missing data. Because missing data due to nonresponse and/or participant attrition is common in survey research and can result in biased results if not attended to, the presence of missing data was explored prior to subsequent analyses. Given the purposes and properties of IRT modeling, it is possible to analyze the dataset in the presence of missing data, using only the completed items in the calculation of person locations along the trait continuum (Boone, Staver, & Yale, 2014). In fact, common practices such as listwise deletion, replacing missing data with a person's mean score across all related items, and replacing missing data with the full sample's mean score across one particular item when utilizing IRT modeling is not recommended (Boone, Staver, & Yale, 2014). When conducting other analyses, however, full datasets are required; therefore, Little's Missing Completely at Random test (MCAR; Little, 1988) was conducted using the BaylorEdPsych package (Beaujean, 2012) in RStudio (RStudio Team, 2018), a free statistical computing and graphics software. Additional explorations of the data were conducted to determine not only the proportion of missing data, but also potential patterns in missing data.

Research Question #1. To what extent, if at all, does the IDCIS represent four distinct constructs, including teachers' attitudes, subjective norms, perceived behavioral control, and behavioral intentions toward IEP data collection?

Confirmatory factor analysis. In answering this question, a separate confirmatory factor analysis (CFA) was completed for each construct (i.e. IDCIS subscales) using the lavaan package (Rosseel, 2012) in RStudio, a free statistical computing and graphics software (RStudio Team, 2018). The CFA for each IDCIS subscale was completed using the fullest dataset available which included 368, 356, 337, and 333 responses for *attitudes*, *subjective norms*, *perceived behavioral control*, and *intentions,* respectively. Given the ordinal nature of the data, the robust correction of the diagonally weighted least squares (DWLS) estimation was chosen, as it does not make any distributional assumptions.

In addition to the commonly reported chi-square goodness-of-fit statistic (χ^2), the standardized root mean square residual (SRMR) and Bentler's Comparative Fit Index (CFI) were examined to determine the strength of each model fit, thus providing validity evidence supporting the extrapolation inference. Because fit statistics are rooted in classical test theory and their recommended criteria for interpretation are based on the traditional Maximum Likelihood (ML) estimation technique involving continuous variables, the use of strict cut-off values to guide model acceptance or rejection is not advocated (e.g., Marsh, Hau, & Wen, 2004), as these conditions are not aligned with the IDCIS and its development. Furthermore, given the issues associated with χ^2, including its hard-to-meet assumptions and deteriorating performance as sample size increases, significant values are common regardless of model fit (Bentler, 1990); therefore, more weight was placed on the supplemental indices when evaluating model fit. As such, Hu and Bentler's recommendation of CFI values "close to" .95 "in combination with a cutoff value close to .09 for SRMR" were loosely applied (1999, p.27).

Following the examination of model-fit indices, factor loadings were examined to determine the extent to which each item contributed to the intended factors. As a preliminary guideline, items with factor loadings below .50 were considered for removal due to less than superior levels of association. Items with low factor loadings relative to all other loadings were also considered for removal.

Finally, a multi-factor CFA including all factors and all retained items corresponding to each factor was conducted using the robust correction of the DWLS estimation technique. A total of 333 responses (i.e., the fullest dataset for *intentions*) were used in the multi-factor CFA. Overall model fit indices (i.e., χ^2, SRMR, and CFI) were reviewed to confirm model fit and factor loadings were examined to ensure all retained items contributed to the intended factor. The main reasoning for conducting this analysis, however, was to examine the correlations between factors.

Research Question #2. Which items serve as quality indicators of each construct, such that valid and reliable inferences about teachers' IEP data collection intentions—regardless of level—can be made using the resulting data?

Item Analysis. Following all CFAs, separate item analyses were conducted on each IDCSI subscale using the free psychometric software, jMetrik (Meyer, 2018) to corroborate the results of the CFA and provide additional information about item and subscale quality. Described as a "procedure for quantifying various characteristics of test items", item analysis produces preliminary data regarding the level of difficulty associated with each item as well as information regarding each item's ability to distinguish between individuals with low and high levels of the construct. In addition to the CFA output, the mean item response for each item and item-total correlations (i.e., item discrimination) were used to support decisions regarding item retention or removal. Items with discrimination values less than 0.30 and that did not contribute to a wider range of difficulty across the subscale were considered for removal (Meyer, 2014). Finally, each subscale's internal consistency was calculated using Guttman's lambda-2 ($\lambda 2$) measure of reliability, with values greater than .80 suggesting the items are similar and will produce consistent scores (Meyer, 2014).

Item response theory modeling. Based on the results of each factor analysis and corresponding item analysis, item response theory (IRT) modeling was completed to evaluate each subscale's ability to produce interval-level data that can be used to make valid and reliable inferences about the IEP data collection intentions of ECSE teachers, irrespective of their level of each construct measured. To evaluate the functioning of all subscales within the IDCIS, the partial credit model (Master, 1982), an extension of the Rasch model, was used to estimate difficulty parameters using jMetrik (Meyer, 2018). This model was chosen for several reasons including (a) the polytomous nature of the data, (b) the relatively small sample size; (c) its ability to allow the comparison of threshold parameters of any two items independent of the sample of respondents; and (d) its ability to allow the comparison of scores of any two respondents regardless of the subset of items administered (Reese & Masse, 2004).

A number of important factors were considered in reviewing the Rasch output; all providing additional validity evidence supporting the extrapolation inference based on item functioning. First, to provide evidence of model fit, weighted mean square (i.e., infit) and unweighted mean square (i.e., outfit) values with respect to both items and persons were examined. While outfit is sensitive to an individual's responses to items with difficulties far from their level of the trait values, infit is more sensitive to an individual's responses to items with difficulty levels that correspond with their level of the trait. Items with infit and outfit values between .50 and 1.50 are considered most productive to measurement (Linacre, 2012); items with values outside of this range received further investigation. Because items are expected to behave better than persons (Linacre, 2012), it is common to be less stringent in the application of rules related to person fit. Additionally, regardless of their pattern of responses, every participant was viewed as an important part of the population for which the IDCIS was created. For these reasons, all potential outliers were retained.

Second, item threshold parameters were examined based on their sequence and spread. The sequencing of thresholds was used to evaluate the functioning of each response scale, while the spread between thresholds was used to evaluate how much information each item produced; although neither was used as grounds for item elimination. Third, Wright maps were created and analyzed to compare the distribution of IDCIS items based on their difficulties to the distribution of persons based on the level of the latent trait they possess (represented by theta, the standard unit of measurement in IRT modeling). Wright maps were used to substantiate the ability of each IDCIS subscale to differentiate between individuals at different levels of each trait, thus providing additional validity evidence based on the internal structure of the scale.

Finally, person reliability and separation statistics were reviewed in order to determine the overall quality of each subscale. Person reliability represents each subscale's ability to accurately order individuals based on their level of latent trait being measured and values at or above 0.80 are preferred (Meyer, 2014). Person separation represents each subscale's ability to *consistently* score and rank individuals and values above 2.0 are preferred (Meyer, 2014).

Research Question #3. Does the TPB (Ajzen, 1985) serve as an appropriate theoretical model to measuring teachers' IEP data collection intentions, such that teachers' attitudes, subjective norms, and perceived behavioral control explain a significant amount of the variance in ECSE teachers' intent to engage in future IEP data collection?

Structural modeling. In answering the last research question, structural modeling—the final stage in structural equation modeling (SEM)—was used to examine 1) the correlation between teachers' attitudes, subjective norms, and perceived behavioral control; and 2) the amount of variance in teachers' intentions that can be explained by their attitudes, subjective norms, and perceived behavioral control. At the conclusion of all other data analysis procedures, the observed variance/covariance matrix based on participants' observed responses to all retained IDCIS items was compared to the implied variance/covariance matrix to determine model fit. In structural modeling, the evaluation of fit is based on the structural model specified by the researcher. Because the construct of perceived behavioral control was separated into two constructs—self-efficacy and controllability—based on the results of the CFA, three structural models were specified and tested (see Figure 9). Model 1 includes *attitudes, subjective norms, self-efficacy,* and *controllability* as exogenous variables and *intentions* as the sole exogenous variable. C*ontrollability* was removed from Model 2, while all other variables remained the same. Finally, Model 3 included *controllability* in place of *self-efficacy.*

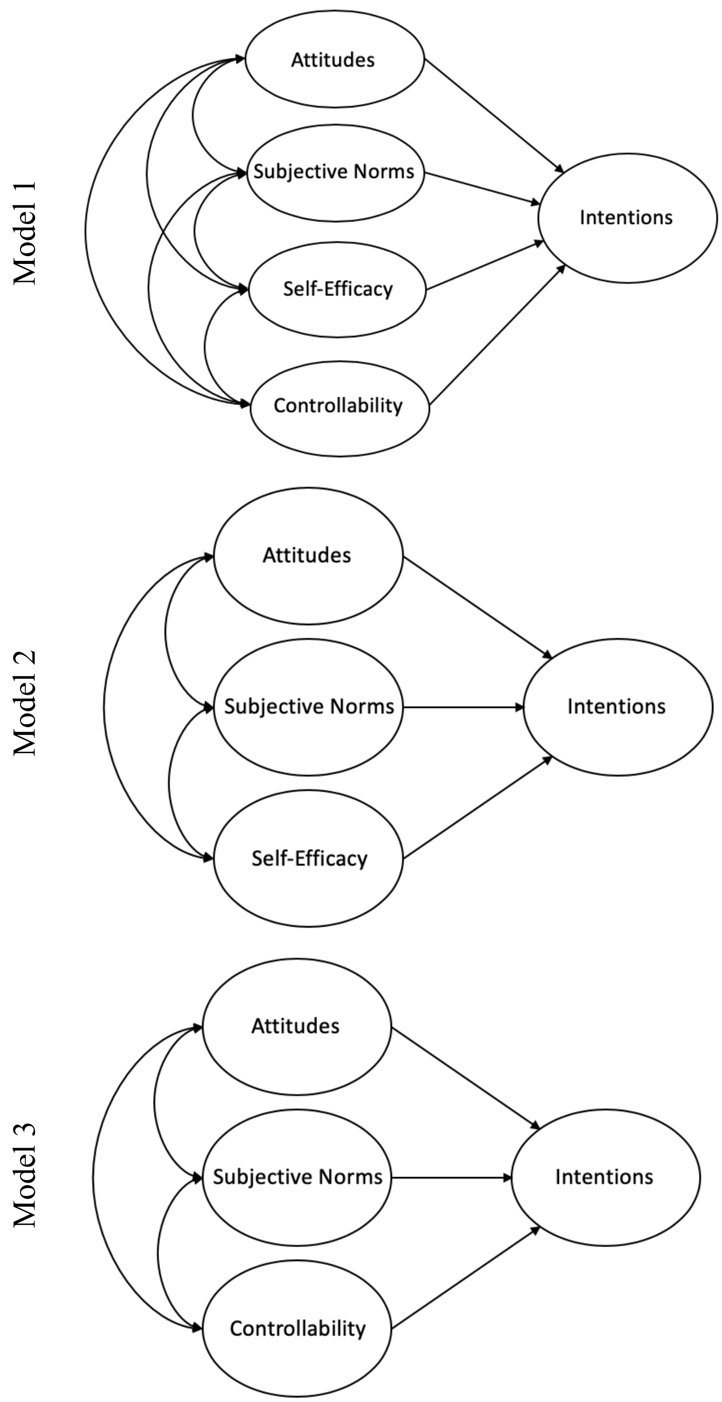

Figure 9. Structural Models Specified and Tested.

Similar to all CFAs completed, the robust correction of the DWLS estimation was chosen given the ordinal nature of the data, and χ^2, CFI, and SRMR were reported to support the evaluation of model fit. Additionally, factor loadings and error terms associated with each retained item, correlation coefficients between exogenous variables, beta weights explaining the impact of each exogenous variable on intentions, and the disturbance term associated with intentions are reported. Finally, the R-squared (R^2) was calculated to determine the proportion of variance in intentions that was explained by all exogenous variables combined. This final stage of data analysis provided additional validity evidence supporting the extrapolation inference described in the IDCIS's IUA—validity evidence based on how each construct is related.

Chapter 4 – Results

Following a summary of missing data, the results are presented in three separate sections, corresponding with each of the three research questions. Each section will be further subdivided into construct-specific segments starting with the results associated with the predictors of intentions—attitudes, subjective norms, and perceived behavioral control—followed by the results associated with intentions.

Analysis of Missing Data

Based on the full dataset consisting of 368 responses, the analysis of missing data revealed that 7.1% of the data were missing. When deconstructed into separate datasets representing each IDCIS subscale, less than 0.01% of the attitude data were missing, 3.2% of the subjective norm data were missing, 8.8% of the perceived behavioral control data were missing, and 9.7% of the intentions data were missing. This pattern of missing data illustrates an increase in participant attrition as they progressed through each section of the survey—only 333 of the 368 respondents maintained participation through the intentions section—suggesting the data included in the full dataset were not missing completely at random (MCAR), a common problem in survey research.

Because subsequent analyses were completed on each separate construct in isolation, an analysis of missing data based solely on the sample of participants who started each section of the survey was completed. Results indicated that less than 0.01% of the data in the attitude sample were missing, 0.2% (based on 356 total responses) of the data in the subjective norms sample were missing, 0.4% (based on 337 total responses) of the data in the perceived behavioral control sample were missing, and 0.2% (based on 333 total responses) of the data in the intentions sample were missing. Based on the results of Little's MCAR test (Little, 1988), the missing data in each of these isolated datasets were found to be MCAR, with the exception of the intentions dataset. Given the negligible amount of missing data in this dataset, however, it was assumed that the data were missing at random (MAR). Based on these analyses, listwise deletion—the default method in the lavaan package—was used both when conducting the CFAs and final structural modeling.

Research Question #1

In determining the extent to which the IDCIS represents four distinct constructs (i.e., teachers' attitudes, subjective norms, perceived behavioral control, and intentions), a confirmatory factor analysis (CFA) was completed with respect to each measured construct. Model fit statistics for each construct's full model and all subsequent model modifications are presented in Table 1. Factor loadings associated with each model are presented in Table 2 through Table 7.

Attitudes. For the full model (Attitudes 1), which included all nine proposed indicators of attitude, results of the chi-square goodness-of-fit test was significant (χ^2 (*df* = 27) = 308.254, $p < 0.01$). Additional fit statistics, however, indicated adequate model fit; the Bentler's Comparative Fit Index (CFI) and standardized root mean square residual (SRMR) were .95 and .07, respectively. All nine indicators showed significant positive factor loadings, with standardized coefficients ranging from .65 to .90 (see Table 2). After further consideration of these nine indicators, item A2 (*How useful, if at all, is IEP data collection?*) was removed due to redundancy; it appeared to capture what was being asked in items A3-A8. For the modified model (Attitudes 2), which included eight indicators of attitude, results of the chi-square goodness-of-fit test was also significant (χ^2

($df = 20$) = 92.63, $p < 0.01$); however, additional fit statistics indicated that the removal of item A2 improved the overall model fit. Superior fit was represented by a CFI and SRMR of .98 and .05, respectively. Factor loadings for the modified model remained high and ranged from .64 to .91 (see Table 2).

Table 1

Confirmatory Factor Analysis Model Fit Statistics

Model Name: IDCIS Items	*n*	χ^2	*df*	*p*	CFI	SRMR
Attitudes 1: A1-A9	365	308.25	27	0.00	.962	.074
Attitudes 2: A1, A3-A9	365	92.63	20	0.00	.988	.047
Subjective Norms 1:SN1-SN8	352	167.96	20	0.00	.936	.095
Subjective Norms 2: SN1, SN2, SN5-SN8	354	74.67	9	0.00	.969	.077
Perceived Behavioral Control 1: SEC1-SEC5, SED1-SED5, PBC1-PBC3, B1-B17	310	3504.79	405	0.00	.528	.164
Perceived Behavioral Control 2: SEC1-SEC5, SED1-SED5, B1-B17	312	3300.75	324	0.00	.537	.174
Self-Efficacy 1: SEC1-SEC5, SED1-SED5	333	678.33	35	0.00	.829	.152
Self-Efficacy 2: SEC1-SEC5	334	177.52	5	0.00	.896	.105
Controllability 1: B1-B17	316	1420.31	119	0.00	.668	.152
Controllability 2: B1, B2, B4-B17	318	1317.90	104	0.00	.644	.146
Intentions 1: I1-I7	328	404.17	14	0.00	.894	.176
Intentions 2: I1, I2, I5-I7	328	181.87	5	0.00	.954	.155

Table 2

Factor Loadings for Attitudes Models

Attitudes 1		Attitudes 2	
Item	Factor Loading	Item	Factor Loading
A1	.870	A1	.750
A2	.860		
A3	.819	A3	.838
A4	.816	A4	.841
A5	.891	A5	.911
A6	.768	A6	.790
A7	.853	A7	.868
A8	.900	A8	.910
A9	.659	A9	.640

Subjective norms. For the full model (Subjective Norms 1), which included all eight proposed indicators of subjective norms, results of the chi-square goodness-of-fit test was significant (χ^2 ($df = 20$) = 167.96, $p < 0.01$). Additional fit statistics also indicated less than adequate model fit; the CFI and SRMR were .91 and .10, respectively. After reviewing the factor loadings in Table 3, items SN3 (*How important, if at all, is IEP data collection to your students' parents?*) and SN4 (*How often, if ever, do you observe your coworkers engaging in IEP data collection?*) were removed due to their low standardized coefficients in relation to all other indicators included in the model. The modified model (Subjective Norms 2), which included only six indicators of subjective norms, resulted in a significant chi-square goodness-of-fit test (χ^2 ($df = 9$) = 74.672, $p < 0.01$). Additional fit statistics, however, indicated that the removal of items SN3 and SN4 improved the overall model fit, resulting in an acceptable fitting model. This superior fit was represented by a CFI and SRMR of .95 and .08, respectively. All six indicators in the modified model showed significant positive factor loadings, with standardized coefficients ranging from .52 to .88 (see Table 3).

Table 3

Factor Loadings for Subjective Norms Models

Subjective Norms 1		Subjective Norms 2	
Item	Factor Loading	Item	Factor Loading
SN1	.518	SN1	.516
SN2	.640	SN2	.543
SN3	.433		
SN4	.320		
SN5	.789	SN5	.801
SN6	.866	SN6	.869
SN7	.875	SN7	.883
SN8	.760	SN8	.768

Perceived behavioral control. For the full model (Perceived Behavioral Control 1), which included all 29 indicators of perceived behavioral control (10 aligned with self-efficacy, and the remaining 19 aligned with controllability), a significant chi-square goodness-of-fit test (χ^2 (df = 405) = 3504.79, $p < 0.01$) and a CFI and SRMR of .50 and .16 respectively, suggested unsatisfactory model fit. Based on the resulting factor loadings presented in Table 4, items PBC1 (*The quality of my students' IEP objectives is entirely up to me.*), PBC2 (*The choice of tools I use to collect IEP data is entirely up to me.*), and PBC3 (*Whether or not I engage in IEP data collection is entirely up to me.*) were removed due to their low standardized coefficients. The modified model (Perceived Behavioral Control 2), which included 26 indicators (10 aligned with self-efficacy, and 16 aligned with controllability) also resulted in a significant chi-square goodness-of-fit test (χ^2 (df = 342) = 3300.75, $p < 0.01$) and the additional indices—CFI of .50 and SRMR of .17—confirmed poor model fit. Based on these results and corroborated by the inconsistent way in which the construct of perceived behavioral control has been measured in the past, it was decided that self-efficacy and controllability be analyzed as separate constructs.

Table 4

Factor Loadings for Perceived Behavioral Control Models

Perceived Behavioral Control 1		Perceived Behavioral Control 2	
Item	Factor Loading	Item	Factor Loading
SEC1	.556	SEC1	.559
SEC2	.629	SEC2	.630
SEC3	.616	SEC3	.611
SEC4	.651	SEC4	.648
SEC5	.705	SEC5	.704
SED1	.462	SED1	.462
SED2	.508	SED2	.511
SED3	.564	SED3	.560
SED4	.684	SED4	.686
SED5	.676	SED5	.676
PBC1	.038	PBC1	
PBC2	.059	PBC2	
PBC3	-.030	PBC3	
B1	.694	B1	.691
B2	.784	B2	.779
B3	.718	B3	.716
B4	.513	B4	.516
B5	.514	B5	.517
B6	.492	B6	.495
B7	.475	B7	.478
B8	.580	B8	.584
B9	.500	B9	.503
B10	.458	B10	.461
B11	.414	B11	.421
B12	.410	B12	.415
B13	.611	B13	.612
B14	.515	B14	.511
B15	.534	B15	.537
B16	.447	B16	.450
B17	.425	B17	.429

Self-efficacy. For the full model (Self-Efficacy 1), which included 10 indicators of self-efficacy (five items representing perceived difficulty and five items representing perceived confidence), a significant chi-square goodness-of-fit test (χ^2 (df = 35) = 678.33, $p < 0.01$), a CFI of .78, and a SRMR of .15 suggested insufficient model fit. Factor loadings, however, were all positive and above .50 (ranging from .59 to .78) as illustrated in Table 5. After reviewing the literature on self-efficacy, it was decided that the five items originally created to represent individuals' perceived confidence (i.e., SEC1-SEC5) best embodied a common conceptualization of self-efficacy and therefore would be used in isolation to evaluate self-efficacy.

Table 5

Factor Loadings for Self-Efficacy Models

Self-Efficacy 1		Self-Efficacy 2	
Item	Factor Loading	Item	Factor Loading
SEC1	.711	SEC1	.724
SEC2	.779	SEC2	.816
SEC3	.753	SEC3	.739
SEC4	.748	SEC4	.797
SEC5	.765	SEC5	.687
SED1	.617		
SED2	.586		
SED3	.707		
SED4	.745		
SED5	.728		

For the modified model (Self-Efficacy 2), which included 5 indicators of self-efficacy (i.e., perceived confidence), the chi-square goodness-of-fit test remained significant (χ^2 (df = 5) = 177.52, $p < 0.01$). While the additional fit statistics—CFI of .79 and SRMR of .11—were an improvement over the previous model, these numbers also suggest less than adequate fit. All five indicators in the modified model, however, showed significant positive factor loadings with standardized coefficients ranging from .69 to .82 (see Table 5). Considering the implications of these fit statistics, this model was used moving forward regardless of the measures of overall fit.

Controllability. The full model (Controllability 1) was comprised of 20 indicators of perceived behavioral control, which included three general items measuring an individual's control over their data collection behavior and 17 specific items assessing the extent to which common barriers to IEP data collection decreases individuals' data collection practices. This model resulted in a significant chi-square goodness-of-fit test (χ^2 ($df = 170$) = 1741.18, $p < 0.01$) and a CFI and SRMR of .58 and .15 respectively, suggesting poor model fit. Based on the resulting factor loadings presented in Table 6, items PBC1, PBC2, and PBC3 were removed due to their low standardized coefficients. The modified model (Controllability 2), which included only the 17 items targeting common barriers, also resulted in a significant chi-square goodness-of-fit test (χ^2 (df = 119) = 1420.31, $p < 0.01$) and higher than preferred CFI and SRMR values (.62 and .15, respectively), suggesting another poor fitting model. Factor loadings ranged from .45 to .83, with only one loading that fell under .50 (see Table 6).

Table 6

Factor Loadings for Controllability Models

Controllability 1		Controllability 2	
Item	Factor Loading	Item	Factor Loading
B1	.704	B1	.658
B2	.830	B2	.743
B3	.736		
B4	.621	B4	.640
B5	.621	B5	.641
B6	.583	B6	.597
B7	.562	B7	.575
B8	.668	B8	.687
B9	.615	B9	.624
B10	.587	B10	.612
B11	.445	B11	.477
B12	.517	B12	.548
B13	.679	B13	.699
B14	.596	B14	.593
B15	.569	B15	.586
B16	.509	B16	.518
B17	.509	B17	.515

After further consideration, item B3 (*How often, if ever, does the availability of time to analyze and interpret IEP data decrease your ability to engage in daily IEP data collection?*) was removed due to the absence of data "use" questions in other subscales of the IDCIS. The adjusted model (Controllability 3), which included 16 items targeting common barriers, resulted in a significant chi-square goodness-of-fit test (χ^2 (df = 104) = 1317.90, p < 0.01) and additional fit statistics confirmed less than adequate model fit (CFI of .59 and SRMR of .15). Factor loadings for the modified model remained fairly consistent, ranging from .48 to .74, with only one loading that fell under .50 (see Table 6). Again, considering the implications of these measures of overall model fit, this model was used moving forward regardless of the resulting fit statistics.

Intentions. For the full model (Intentions 1), which included all seven proposed indicators of intentions, the results of the chi-square goodness-of-fit test was significant (χ^2 (df = 14) = 404.17, $p < 0.01$). Additional fit statistics also indicated less than adequate model fit; the CFI and SRMR were .84 and .18, respectively. After reviewing the factor loadings in Table 7, items I3 (*Over the next month, to what extent, if at all, do you intend to engage in IEP data collection by reflecting on your memory of students' skills/behaviors while writing down informal notes?*) and I4 (*Over the next month, to what extent, if at all, do you intend to engage in IEP data collection by talking with others about their observations of students' skills/behaviors while taking informal notes?*) were removed due to their low standardized coefficients. The modified model (Intentions 2), which included the remaining five indicators of intentions, resulted in a non-significant chi-square goodness-of-fit test (χ^2 (df = 5) = 181.87, $p = 0.00$). Additional fit statistics including a CFI of .91 and SRMR of .16 confirmed less than adequate model fit. All remaining indicators (i.e., I1, I2, 15-I7) showed significant positive factor loadings, however, with standardized coefficients ranging from .50 to .94 (see Table 7). This model was used moving forward regardless of the resulting fit statistics.

Table 7

Factor Loadings for Intentions Models

Intentions 1		Intentions 2	
Item	Factor Loading	Item	Factor Loading
I1	.865	I1	.869
I2	.864	I2	.870
I3	.148		
I4	.227		
I5	.534	I5	.504
I6	.936	I6	.935
I7	.885	I7	.888

Multi-factor CFA. Results of the multi-factor CFA revealed a significant chi-square goodness-of-fit test (χ^2 (df = 730) = 2047.828, p = 0.00) and an SRMR and CFI of .10 and .87, respectively. The correlation between factors ranged from .125 to .507, with the strongest correlation between attitudes and intentions and the weakest correlation between subjective norms and self-efficacy. See Table 8 for the correlations among all factors.

Table 8

Correlations Among Factors

	ATT	SN	SE	CON
SN	.484			
SE	.286	.125		
CON	.153	.246	.300	
INT	.507	.355	.331	.217

Note: ATT = attitudes; SN = subjective norms; SE = self-efficacy; CON = controllability; INT = intentions

Research Question #2

In determining which items serve as quality indicators of each construct, such that the IDCIS produces valid and reliable measures of teachers' attitudes, subjective norms, perceived behavioral control, and intentions, item analyses and Rasch modeling were completed on each IDCIS subscale.

Conditions necessary for Rasch modeling that were tested throughout the analyses include the following: 1) unidimensionality of measured constructs, that is a person's level of the construct being measured is the dominant factor contributing to the person's response to each item; 2) local independence of items, meaning a person's response to one item does not impact their response to another item; and 3) model appropriateness (Henard, 2000). The first two assumptions were evaluated via CFAs and item analyses as well as by a final review of IDCIS items. The final assumption was met by ensuring the Rasch partial credit model was compatible with the sample size and type of data (i.e., ordered polytomous responses) and by examining the data to ensure they fit the model through item fit indices provided through jMetrik (Meyers, 2018).

Attitudes. As shown in Table 9, item analysis resulted in discrimination values ranging from .60 to .77, suggesting all items were able to discriminate between individuals with similar levels of attitude. The mean item response ranged from 2.67 to 3.62; item A9 (*To what extent, if at all, do you agree that thinking about IEP data collection gives you a positive feeling?*) was the most difficult to endorse, while item A4 (*To what extent, if at all, do you agree that IEP data collection improves the quality of your progress reporting?*) was the easiest to endorse. Guttman's $\lambda 2$ internal consistency measure of reliability was .89 suggesting that items A1 and A3 through A9 on the Attitudes subscale of the IDCIS hang well together.

Table 9

Item Analysis Results for Attitudes Subscale

Item	Mean	Discrimination
A1	3.56	0.62
A3	3.40	0.69
A4	3.62	0.65
A5	3.25	0.77
A6	3.36	0.60
A7	3.36	0.71
A8	3.33	0.74
A9	2.67	0.53

Results of the Rasch analysis of the final eight items in the Attitude subscale (i.e., items A1 and A3-A9) are presented in Table 10. Item infit ranged from .74 to 1.52 (item A9), while item outfit ranged from .65 to 1.55 (item A9), suggesting good fit. Though the fit statistics for item A9 were slightly above the preferred cut-off, this item was retained given it was designed to sit at the highest level of the construct (i.e., to be endorsed only by those with the highest level of attitude). Person-level infit values ranged from 0.08 to 5.32, while person outfit values ranged from 0.05 to 8.18. A total of 31 individuals (8.4% of the sample) had infit values greater than 2.0, while 24 (6.5% of the sample) had outfit values greater than 2.0. Person reliability and separation values were .86 and 2.48, respectively, suggesting this subscale can be used to accurately order individuals along the subjective norms continuum and consistently score and rank individuals based on their level of subjective norms.

Table 10

Rasch Analysis Results for Attitudes Subscale

Item	Difficulty	Std. Error	WMS	UMS
A1	-1.12	0.13	1.09	1.08
A3	-0.29	0.13	0.92	0.84
A4	-1.44	0.14	0.94	0.83
A5	0.47	0.12	0.74	0.68
A6	-0.59	0.13	1.08	1.04
A7	-0.04	0.13	0.88	0.86
A8	0.07	0.12	0.75	0.65
A9	2.94	0.11	1.52	1.55

Note. Std. Error = standard error of difficulty;
WMS = weighted mean square fit statistic (infit);
UMS = unweighted mean square fit statistic (outfit)

In regard to item threshold parameters, thresholds across all items were appropriately sequenced suggesting the response scale was functioning properly; lower response categories were easier to endorse than higher response categories. Review of the item characteristic curves highlighted in Figure 10 confirmed proper scale functioning and revealed that all four response categories were helpful to the measurement of teachers' attitudes. Additionally, the spread of thresholds ranged from 6.53 logits (item A4) to 8.89 logits (item A6; *To what extent, if at all, do you agree that IEP data collection improves your accountability to others?*), suggesting each retained item provided a large amount of information related to an individual's attitude toward IEP data collection. The Wright map for the Attitude subscale, which compares the range of the construct represented by the items to the range of theta possessed by the individuals in the sample, suggests that the items do a good job of covering all levels of the construct (see Figure 11), though most items appeared to tap individuals at or below 5 logits, while only one item tapped those above 5 logits.

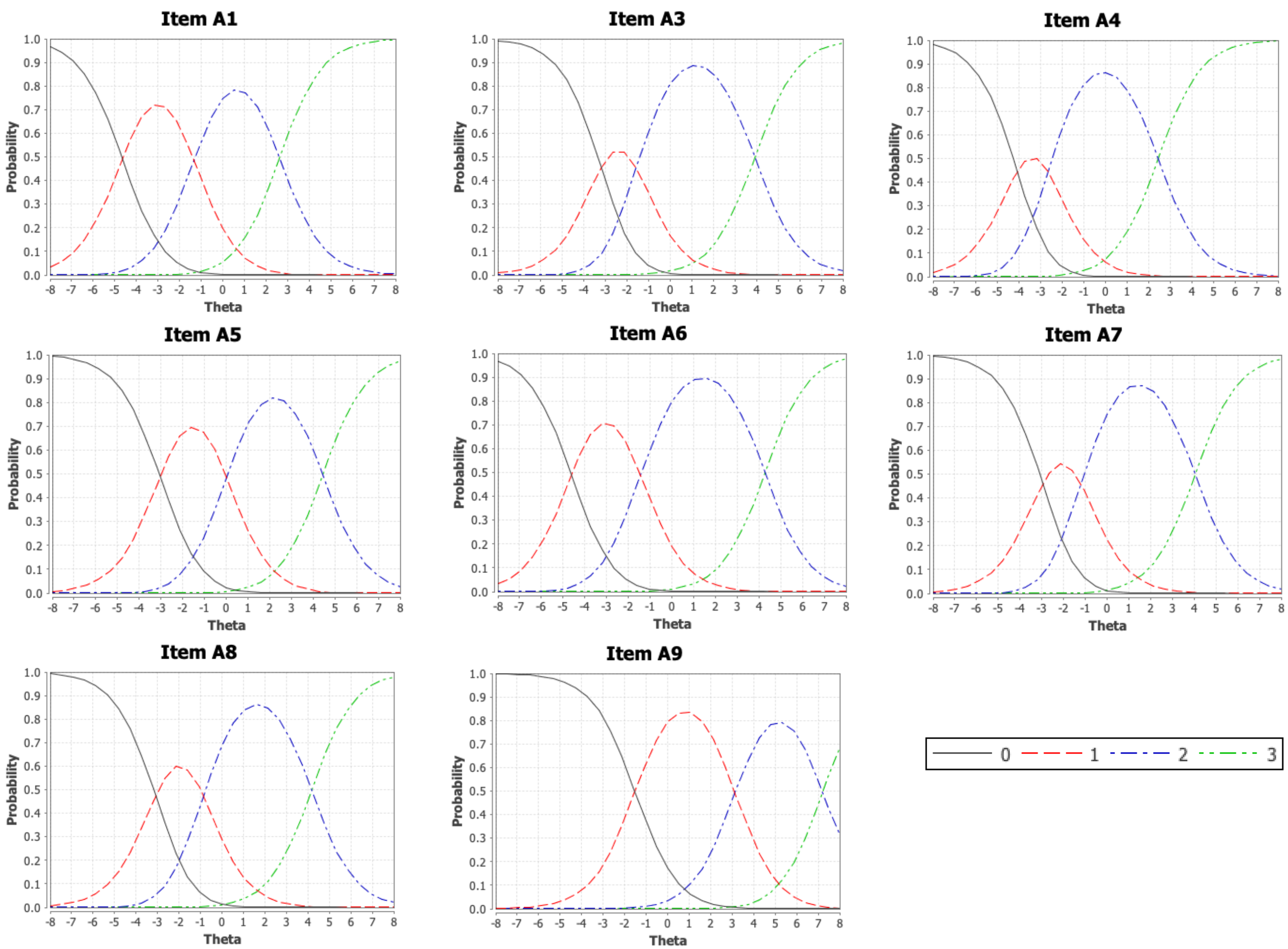

Figure 10. Item Charactieristic Curves for the Attitudes Subscale

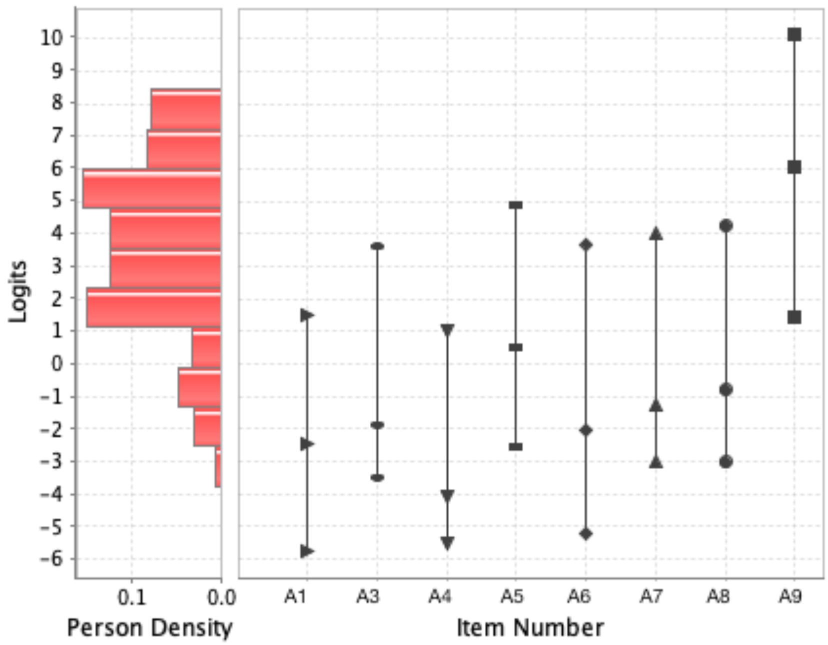

Figure 11. Wright map for the Attitudes Subscale.

Subjective Norms. The results of the item analysis are displayed in Table 11. Discrimination values ranging from .54 to .73 suggest that all items were able to discriminate between individuals with similar levels of subjective norms. The mean item response ranged from 1.72 to 3.11; item SN7 (*How often, if ever, do you observe someone in an ECSE leadership role acknowledge you for your IEP data collection efforts?*) was the most difficult to endorse, while item SN1 (*How important, if at all, is IEP data collection to ECSE leadership in your district?*) was the easiest to endorse. Guttman's $\lambda 2$ was 0.87, suggesting that items SN1, SN2, and SN5-SN8 on the Subjective Norms subscale of the IDCIS hang well together.

Table 11

Item Analysis Results for Subjective Norms Subscale

Item	Mean	Discrimination
SN1	3.56	0.62
SN2	3.40	0.69
SN5	3.62	0.65
SN6	3.25	0.77
SN7	3.36	0.60
SN8	3.36	0.71

Results of the Rasch analysis of the final six items in the Subjective Norms subscale (i.e., items SN1, SN2, and SN5-SN8) are presented in Table 12. All items had infit values between .78 and 1.23 and item outfit values ranged from .77 to 1.44; both suggesting good fit. Person-level infit ranged from .05 to 6.5, while person outfit ranged from .05 to 13.3. A total of 25 individuals (7% of the sample) had infit values greater than 2.0, while 32 (9% of the sample) had outfit values greater than 2.0. Person reliability was 0.80 and person separation was 1.99, both situated at or near the lowest values in the acceptable range.

Table 12

Rasch Analysis Results for Subjective Norms Subscale

Item	Difficulty	Std. Error	WMS	UMS
SN1	-2.03	0.09	1.17	1.44
SN2	-1.59	0.09	1.23	1.28
SN5	0.27	0.08	0.78	0.77
SN6	1.23	0.08	0.86	1.00
SN7	1.48	0.09	0.84	0.79
SN8	0.62	0.08	0.95	0.94

Note. Std. Error = standard error of difficulty;
WMS = weighted mean square fit statistic (infit);
UMS = unweighted mean square fit statistic (outfit)

Regarding item threshold parameters, lower categories were easier to endorse than higher categories (i.e., thresholds across all items were appropriately sequenced), suggesting the rating scales were functioning properly. Review of the item characteristic curves for items SN5 (*How often, if ever, do you observe someone in an ECSE leadership role communicate with you about IEP data collection?*), SN6 (*How often, if ever, do you observe someone in an ECSE leadership role look at your IEP data?*), and SN7 (*How often, if ever, do you observe someone in an ECSE leadership role acknowledge you for your IEP data collection efforts?*) shown in Figure 12, however, revealed that the "yearly" response option—which was coded 2—added little value to the scale. Response scales for all other items appeared to perform well. Additionally, the spread of thresholds ranged from 1.84 logits (item SN6) to 5.87 logits (item SN2; *How important, if at all, is IEP data collection to your coworkers?*). While most items only spanned 3 or fewer logits, SN1 and SN2 had a much wider range of thresholds, thus tapping a wider range of theta. Based on analysis of the Wright map shown in Figure 13, the final eight items included in the Subjective Norms subscale do a good job of covering all levels of the trait represented in the sample; however, a higher frequency of items are located at the extremes, whereas a higher frequency of individuals fall at the mid-range of theta.

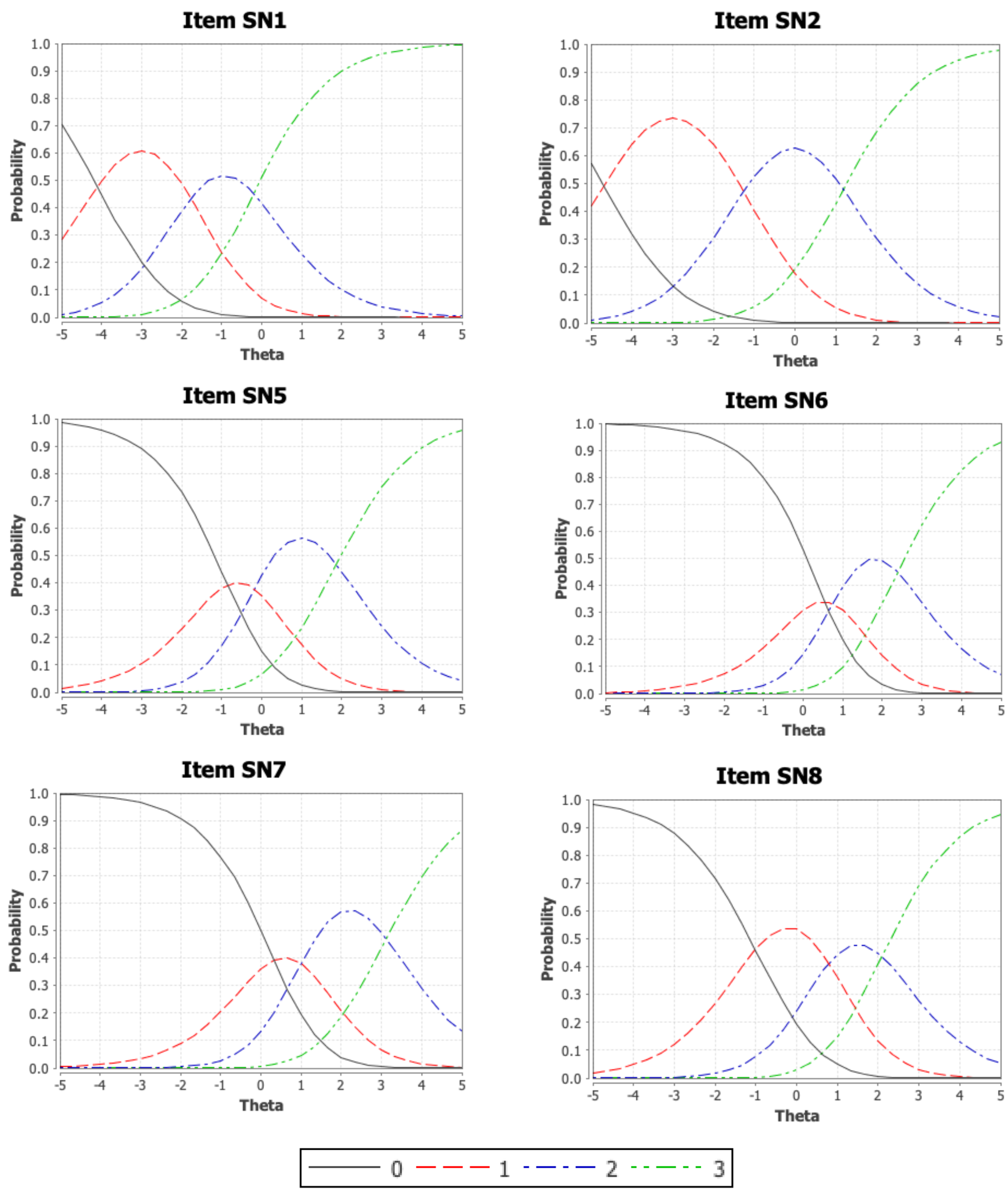

Figure 12. Item Charactieristic Curves for the Subjective Norms Subscale

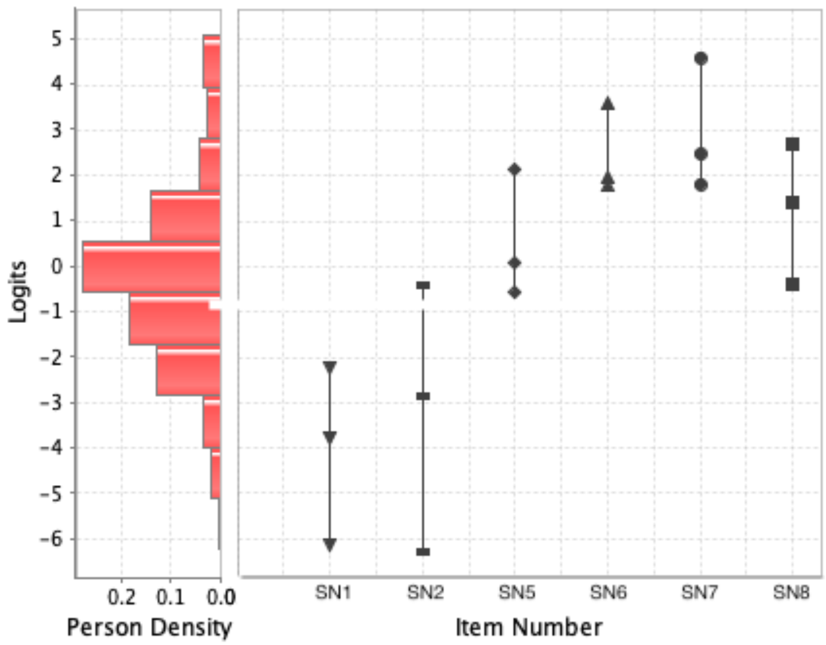

Figure 13. Wright map for the Subjective Norms Subscale.

Perceived Behavioral Control. Based on the results of the confirmatory factor analyses, self-efficacy and controllability underwent separate item and Rasch analyses.

Self-efficacy. Results of the item analysis (see Table 13) supported the inclusion of items SEC1-SEC5 as indicators of self-efficacy. Discrimination values ranged from .64 to .76, suggesting all items were able to discriminate between individuals with similar levels of self-efficacy. The mean item response ranged from 1.29 to 1.74; item SEC5 (*How confident, if at all, are you in your ability to consistently carry out your plan, such that IEP data are collected on a daily basis?*) was the most difficult to endorse, while item SEC1 (*How confident, if at all, are you in your ability to write clear and measurable IEP objectives across all developmental domains?*) was the easiest to endorse. Guttman's $\lambda 2$ internal consistency measure of reliability was 0.93 suggesting that items SEC1-SEC5 on the Self-Efficacy subscale of the IDCIS hang well together.

Table 13

Item Analysis Results for Self-Efficacy and Controllability Subscales

Self-Efficacy			Controllability		
Item	Mean	Discrimin.	Item	Mean	Discrimin.
SEC1	2.00	0.70	B1	2.00	0.70
SEC2	1.91	0.71	B2	1.91	0.71
SEC3	2.74	0.69	B4	2.74	0.69
SEC4	2.52	0.66	B5	2.52	0.66
SEC5	2.39	0.72	B6	2.39	0.72
			B7	2.57	0.76
			B8	2.36	0.80
			B9	2.45	0.76
			B10	2.49	0.77
			B11	2.84	0.78
			B12	2.52	0.76
			B13	2.35	0.80
			B14	2.02	0.72
			B15	2.27	0.74
			B16	2.36	0.74
			B17	2.35	0.76

Note. Discrimin = discrimination

As displayed in Table 14, the Rasch analysis of the final five items in the Self-Efficacy subscale (i.e., items SEC1-SEC5) resulted in infit values between .89 and 1.20 and outfit values ranging from 0.86 to 1.19, both suggesting that all retained items fit well. Person-level infit ranged from .06 to 7.02 and person outfit values were between .06 and 6.80. A total of 45 individuals (13.5% of the sample) had infit values greater than 2.0, while 44 (13.2% of the sample) had outfit values greater than 2.0. Finally, person reliability was 0.81 and person separation was 2.08, highlighting the subscale's ability to accurately order individuals along the self-efficacy continuum and to consistently score and rank individuals based on their level of self-efficacy.

Table 14

Rasch Analysis Results for Self-Efficacy and Controllability Subscales

Self-Efficacy					Controllability				
Item	Difficulty	Std. Error	WMS	UMS	Item	Difficulty	Std. Error	WMS	UMS
SEC1	-1.18	0.12	0.97	0.96	B1	1.01	0.08	1.06	1.07
SEC2	-0.09	0.11	0.89	0.88	B2	1.33	0.08	0.96	0.96
SEC3	0.45	0.10	1.02	1.01	B4	-0.65	0.07	1.29	1.30
SEC4	0.02	0.10	0.89	0.86	B5	-0.28	0.07	1.28	1.30
SEC5	0.80	0.10	1.20	1.19	B6	-0.10	0.07	0.99	0.96
					B7	-0.43	0.08	1.01	1.00
					B8	0.08	0.08	0.83	0.82
					B9	-0.15	0.08	0.92	0.91
					B10	-0.27	0.08	0.94	0.93
					B11	-1.38	0.09	1.07	1.05
					B12	-0.31	0.08	1.02	1.03
					B13	0.02	0.08	0.82	0.81
					B14	0.83	0.08	0.94	0.94
					B15	0.13	0.08	0.91	0.93
					B16	0.05	0.07	1.03	1.13
					B17	0.12	0.08	0.98	0.98

Note. Std. Error = standard error of difficulty; WMS = weighted mean square fit statistic (infit); UMS = unweighted mean square fit statistic (outfit)

In terms of threshold parameters, thresholds across all items were appropriately sequenced suggesting the rating scales were functioning properly. The quality of the rating scale was confirmed by reviewing the item characteristic curve for each item (see Figure 14); all response options were valuable in the measurement of teachers' self-efficacy. All items appeared to cover a wide range of ability levels, with item thresholds ranging from 10.41 logits (item SEC5) to 13.86 logits (item SEC2; *How confident, if at all, are you in your ability to modify existing assessment tools to meet your IEP data collection needs?*). Based on analysis of the Wright map shown in Figure 15, the five items included in the Self-Efficacy subscale do a good job of covering all levels of the trait represented in the sample, with the exception of a small percentage of individuals with the highest levels of theta.

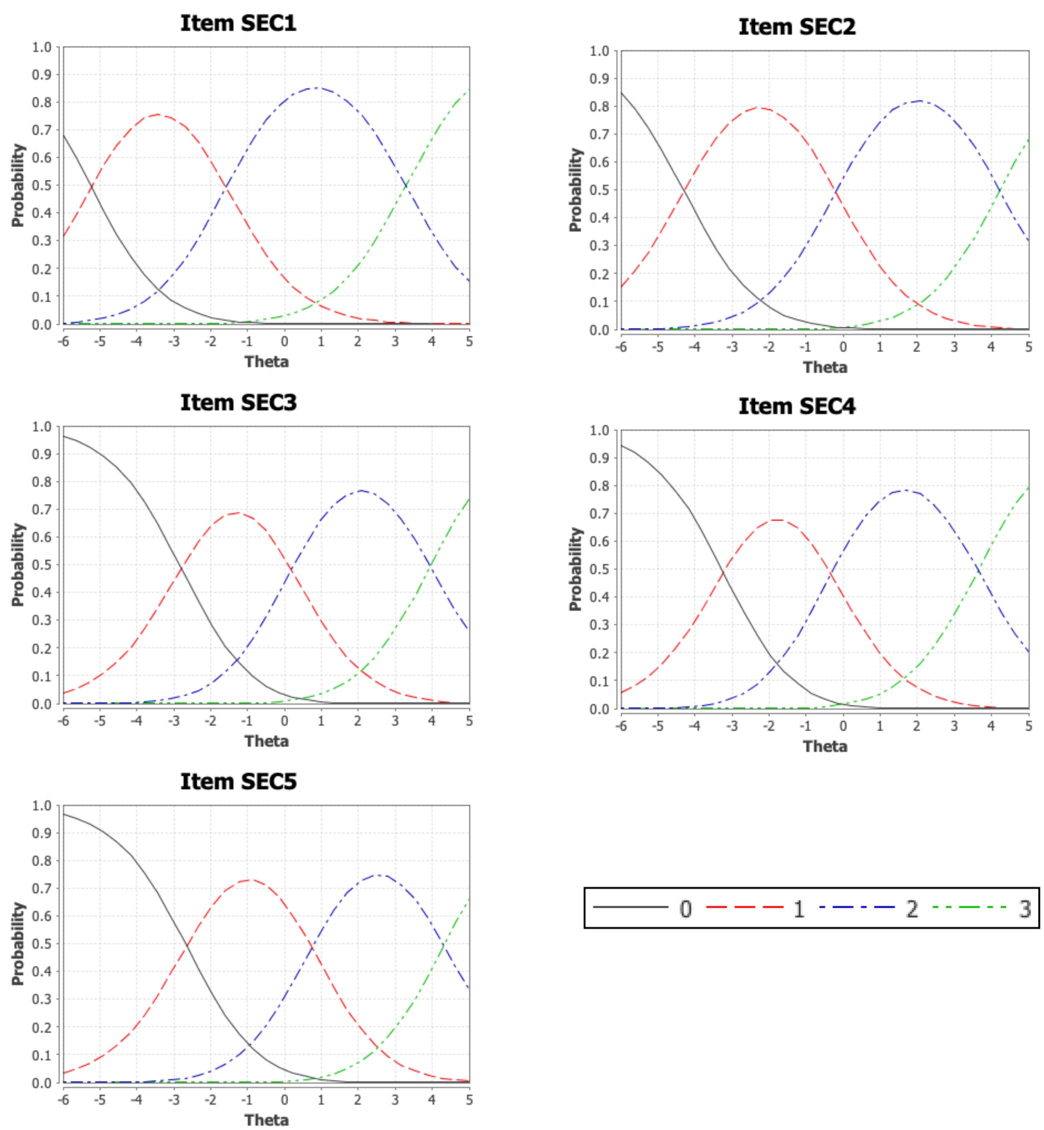

Figure 14. Item Charactieristic Curves for the Self-Efficacy Subscale

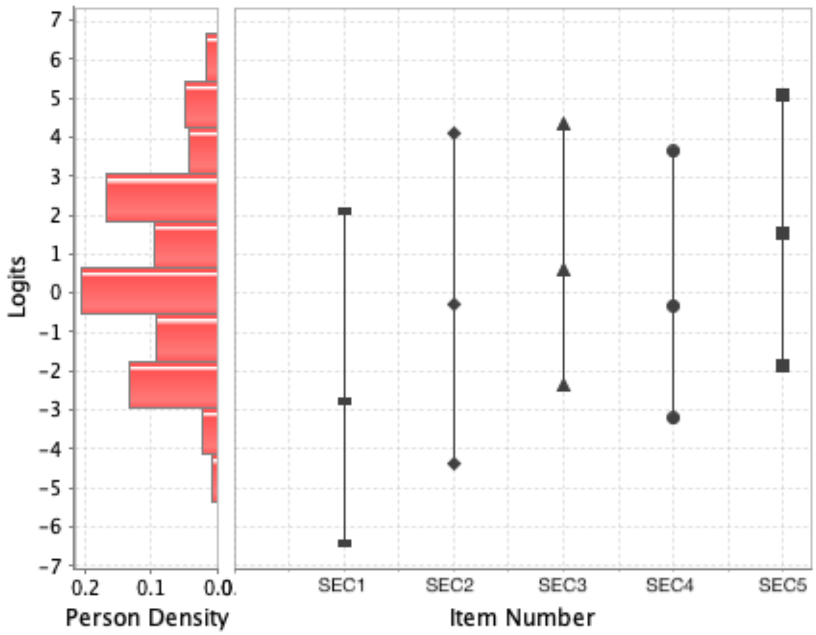

Figure 15. Wright map for the Self-Efficacy Subscale.

Controllability. As displayed in Table 13, the results of the item analysis supported the inclusion of items B1, B2, and B4-B17 as indicators of controllability. Discrimination values ranged from 0.65 to 0.80, suggesting all items were able to discriminate between individuals with similar levels of controllability. The mean item response ranged from 1.91 to 2.85; item B2 (*How often, if ever, does the availability of time to collect IEP data decrease your ability to engage in daily IEP data collection?*) was the most difficult to endorse, while item B11 (*How often, if ever, does the clarity of IEP objectives on your students' IEPs decrease your ability to engage in daily IEP data collection?*) was the easiest to endorse. Guttman's $\lambda 2$ internal consistency measure of reliability was 0.96, suggesting that these items on the Controllability subscale of the IDCIS hang well together.

Results of the Rasch analysis of the final 16 items in the Controllability subscale (i.e., items B1, B2, and B4-B17) are also included in Table 14. Infit values between 0.82

and 1.29 and outfit values ranging from 0.81 to 1.30 suggest proper fit of all items. Person-level infit ranged from 0.04 and 3.09 and person outfit values were between 0.02 and 3.64. A total of 26 individuals (8.2% of the sample) had infit values greater than 2.0, while 21 (6.6% of the sample) had outfit values greater than 2.0. Person reliability was 0.87 and person separation was 2.44, both highlighting the subscale's ability to accurately order individuals along the controllability continuum and consistently score and rank individuals based on their level of controllability.

In regard to item threshold parameters, thresholds across all items were appropriately sequenced suggesting the rating scales were functioning properly; indicating that a barrier "always" decreased one's ability to collect IEP data on a daily basis was easier than indicating that the same barrier "never" decreased data collection abilities. The quality of the rating scale was confirmed by reviewing the item characteristic curve for each item (see Figure 16) and suggested that all response categories were useful to the measurement of teachers' controllability. Compared to other subscales, most items serving as indicators of controllability only covered a narrow range of ability levels, with thresholds ranging from 2.21 (B4) to 6.07 logits (B11; *How often, if ever, does access to commercially-available paper-pencil tools that meet your IEP data collection needs decrease your ability to engage in daily IEP data collection?*). Based on analysis of the Wright map shown in Figure 17, however, the 16 items included in the Controllability subscale do a good job of covering all levels of the trait represented in the sample.

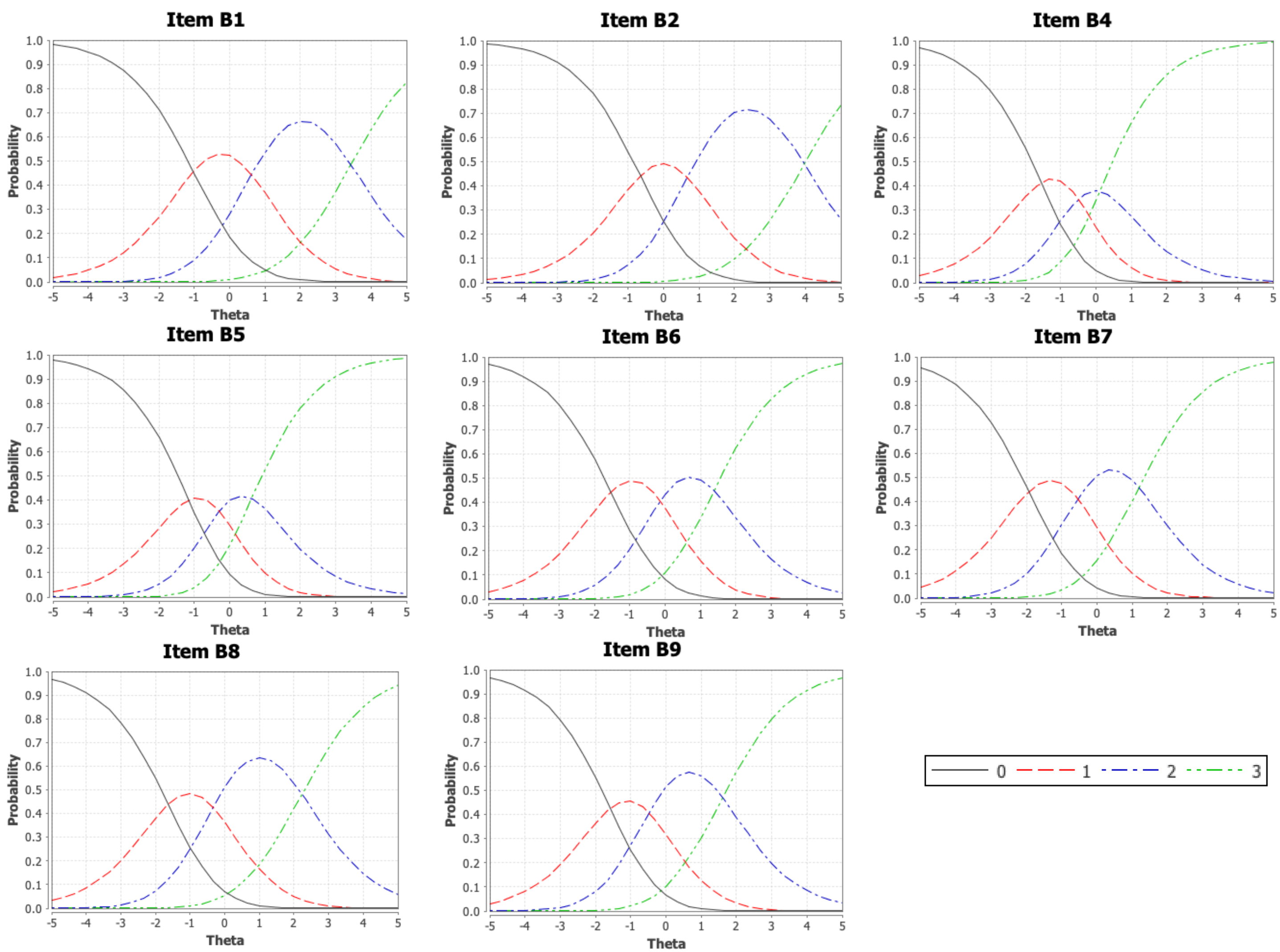

Figure 16. Item Charactieristic Curves for Items B1, B2, and B4-B9 on the Controllability Subscale

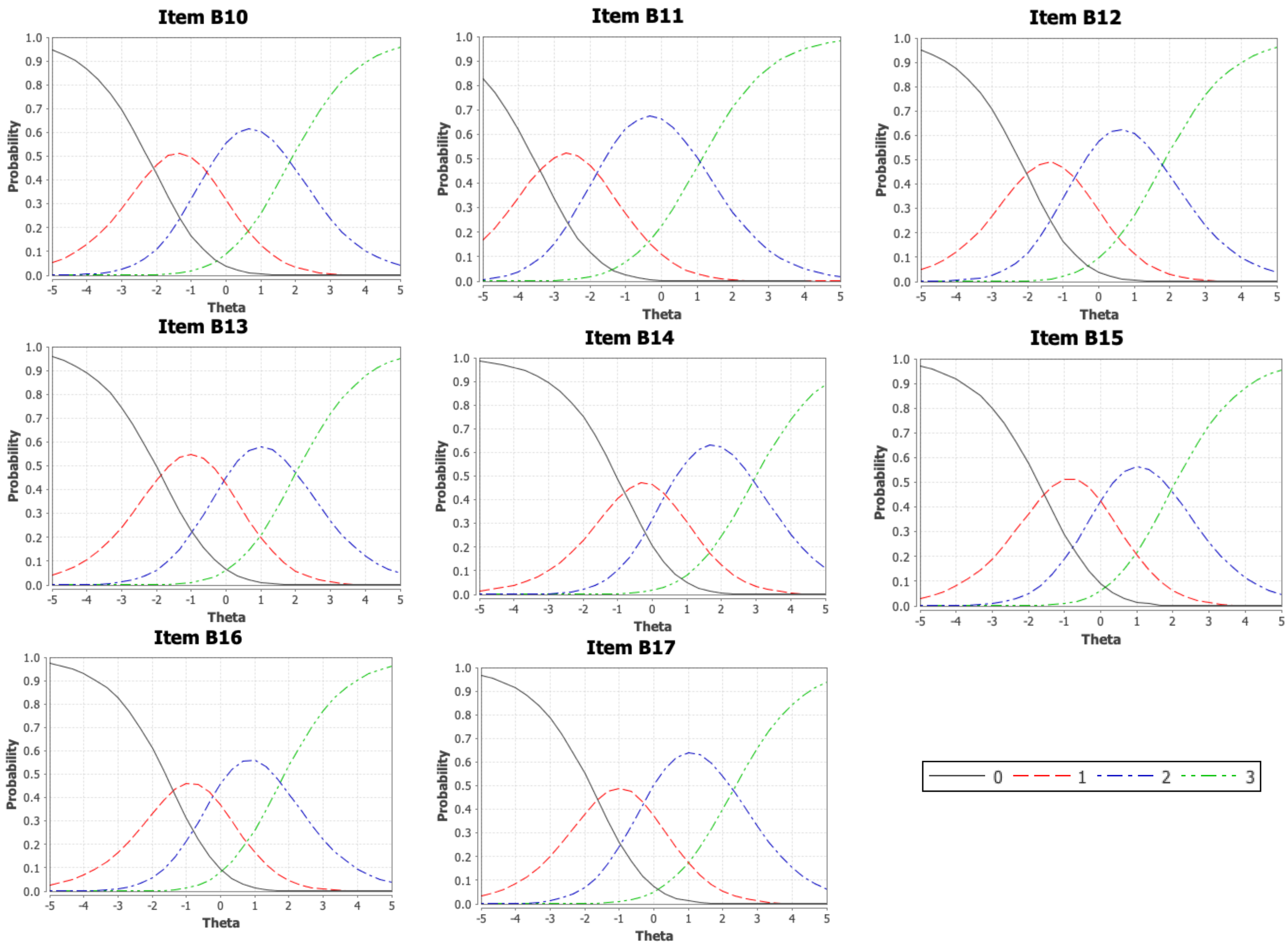

Figure 17. Item Charactieristic Curves for Items B10-B17 on the Controllability Subscale.

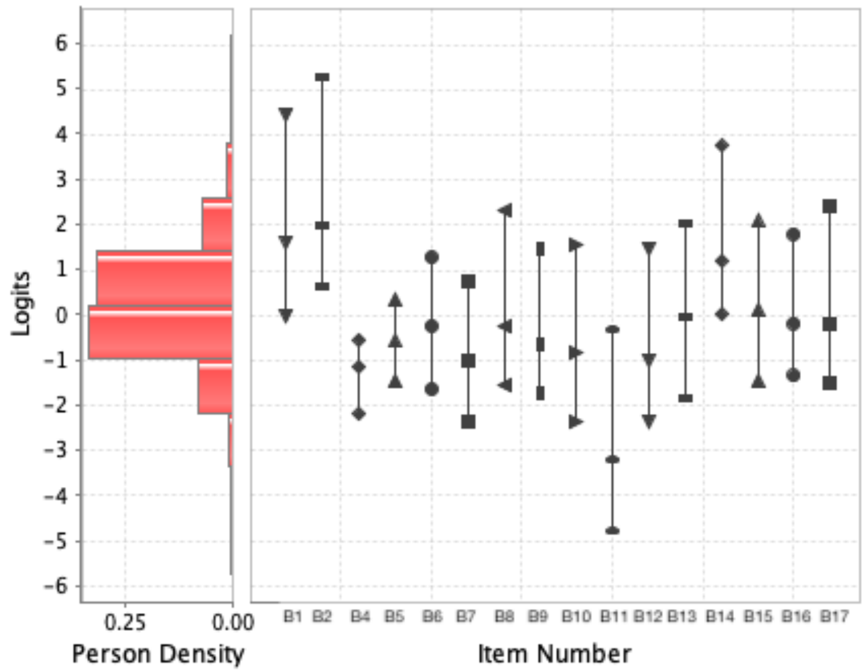

Figure 18. Wright map for the Controllability Subscale.

Intentions. Item analysis resulted in discrimination values ranging from .74 to 0.83 as shown in Table 15, suggesting all retained items were able to discriminate between individuals with similar levels of intentions. The mean item response ranged from 1.68 to 2.18; item I7 (*Over the next month, to what extent if at all, do you intend to observe students while formally documenting the frequency and/or duration of their skills/behaviors AND the type and/or frequency of adult prompts provided?*) was the most difficult to endorse, while item I5 (*Over the next month, to what extent if at all, do you intend to observe students while taking informal notes about their skills?*) was the easiest to endorse. Guttman's $\lambda2$ internal consistency measure of reliability was .92, suggesting that items I1, I2, and I5-I7 on the Intentions subscale of the IDCIS hang well together.

Table 15

Item Analysis Results for Intentions Subscale

Item	Mean	Discrimination
I1	2.16	0.82
I2	2.08	0.78
I5	2.18	0.74
I6	1.82	0.83
I7	1.68	0.77

Results of the Rasch analysis of the final five items in the Intentions subscale (i.e., items I1, I2, and I5-I7) are presented in Table 16. Item infit ranged from .69 to 1.37 and item outfit ranged from 0.68 to 1.54, both suggesting good fit. Person-level infit ranged from 0.10 to 4.57, while outfit ranged from .07 to 5.17. A total of 20 individuals (6.1% of the sample) had infit values greater than 2.0, while 33 (10% of the sample) had outfit values greater than 2.0. Person reliability was .77, while person separation was 1.81, both raising some concerns regarding the overall quality of the subscale.

Table 16

Rasch Analysis Results for Intentions Subscale

Item	Difficulty	Std. Error	WMS	UMS
I1	-1.13	0.13	1.01	1.03
I2	-0.20	0.12	1.08	1.14
I5	-1.15	0.12	1.37	1.54
I6	0.93	0.12	0.69	0.68
I7	1.55	0.10	0.84	0.83

Note. Std. Error = standard error of difficulty;
WMS = weighted mean square fit statistic (infit);
UMS = unweighted mean square fit statistic (outfit)

Regarding item threshold parameters, thresholds across all items were appropriately sequenced suggesting the rating scales were functioning properly. Review of the item characteristic curves highlighted in Figure 18 confirmed proper scale functioning. Additionally, the spread of thresholds ranged from 5.68 logits (item I7) to 7.69 logits (item I1; *To what extent, if at all, do you plan to collect IEP data across all students on your caseload over the next month?*), suggesting each retained item provides a substantial amount of information regarding an individual's intentions to engage in IEP data collection. The Wright map for the Intentions subscale illustrated in Figure 19 suggests that the items do a good job of covering all levels of the construct.

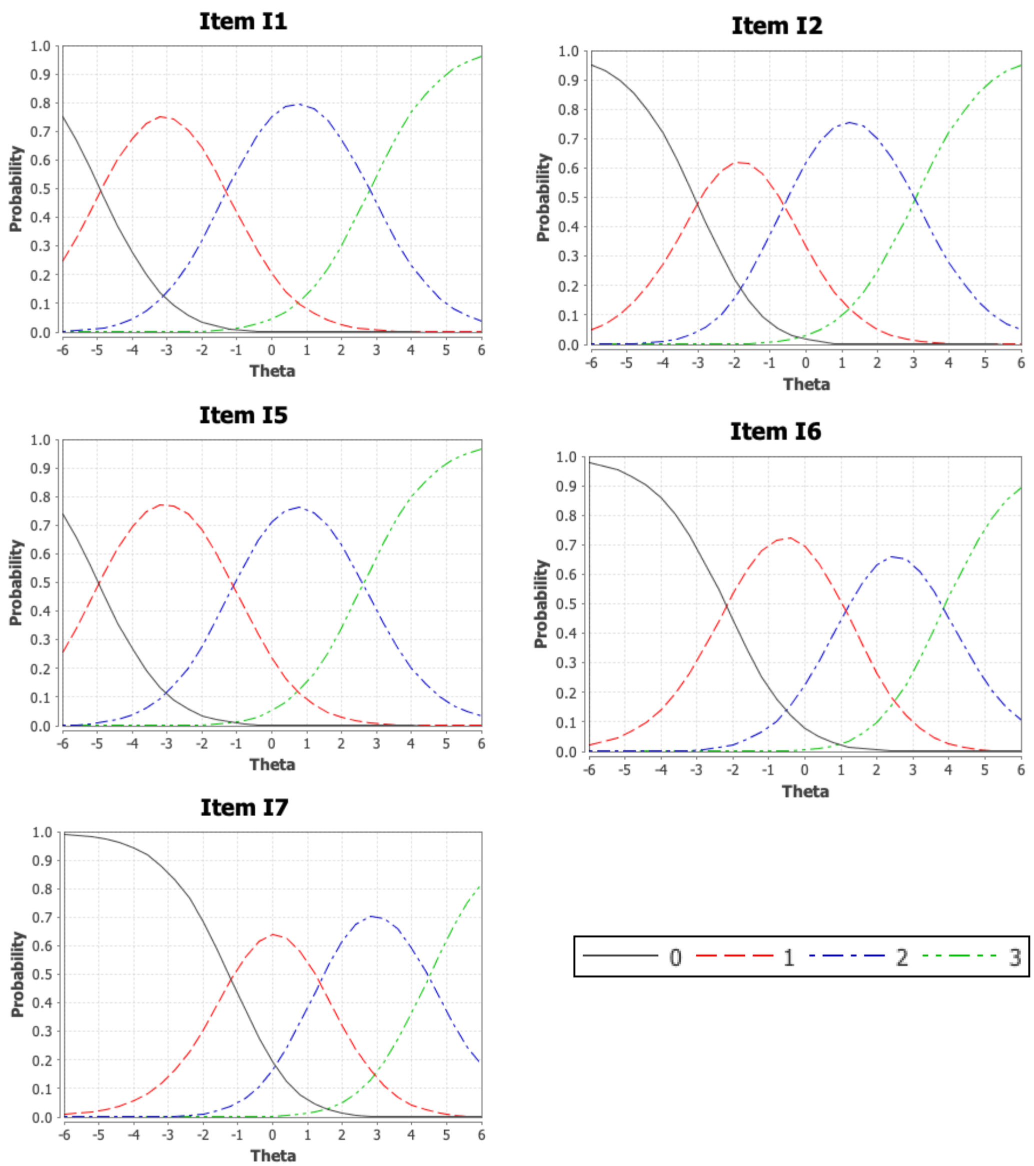

Figure 19. Item Charactieristic Curves for the Intentions Subscale

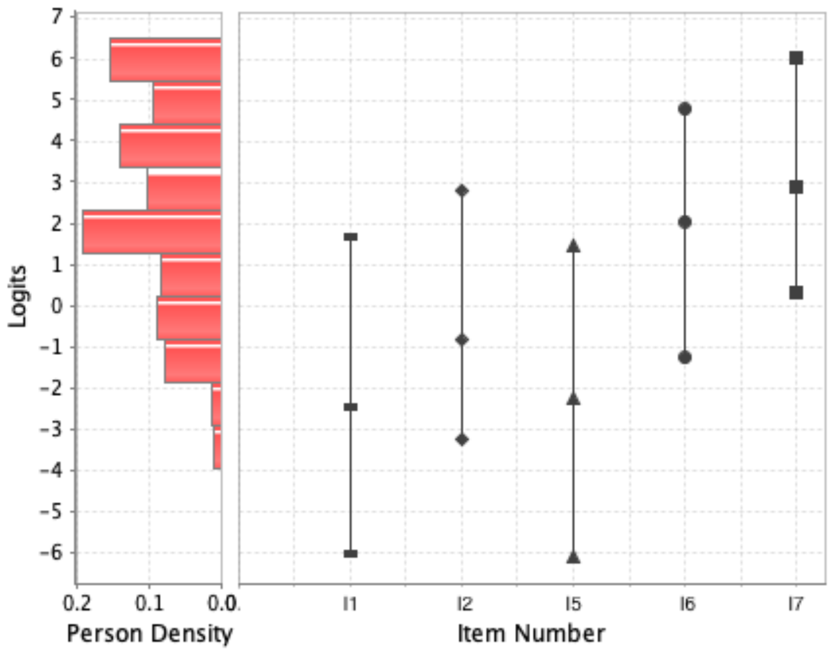

Figure 20. Wright map for the Intentions Subscale.

Research Question #3

In determining the degree to which each IDCIS construct is correlated and how much variance in intentions can be attributed to the predictor variables, suggesting whether the TPB (Ajzen, 1985) serves as an appropriate theoretical model to the measurement of ECSE teacher's IEP data collection intentions, three structural models were tested (see Figure 9). A total of 321 observed responses to the retained IDCIS items were used in the analysis of each model. Based on the commonly reported sample size requirement for SEM—between 5 and 10 cases per parameter estimated (Bentler & Chou, 1987)—321 cases was a sufficient sample size for testing all models. Model 1 estimated 49 parameters resulting in 6.6 cases per parameter, Model 2 estimated 30 parameters resulting in 10.7 cases per parameter, and Model 3 estimated 41 parameters resulting in 7.8 cases per parameter.

As shown in Table 17, Model 2 resulted in the best fit (see Table 18 for the full structural modeling results). While the chi-square test statistic was significant (χ^2 = 943.15 (*df* = 246), *p* = 0.00), the CFI was .924 and the SRMR was .100, suggesting satisfactory fit. The relationships among each set of latent variables are presented in Figure 20. The correlation coefficients representing the relationship between each predictor variable (i.e., attitudes, subjective norms, and self-efficacy) ranged from .11 to .50. The strength of correlations varied, ranging from weak to strong. Standardized regression weights show that attitudes was the strongest predictor of intentions (β = .38, *p* = 0.00), followed by self-efficacy (β = .20, *p* = 0.00) and then subjective norms (β = .16, *p* = 0.004). Together, attitudes, subjective norms, and self-efficacy accounted for approximately 32% of the variance in teachers' intentions, suggesting that the TPB (Ajzen, 1985) is an appropriate theoretical model in organizing preliminary examinations of the many factors impacting ECSE teachers' intentions to engage in IEP data collection.

Table 17

Structural Modeling Results – Overall Model Fit

Model Name: Exogenous Variables	*n*	χ^2	*df*	*p*	CFI	SRMR
Model 1: ATT, SN, SE, CON	302	2047.83	730	0.00	0.870	0.101
Model 2: ATT, SN, SE	319	943.15	246	0.00	0.924	0.100
Model 3: ATT, SN, CON	305	1737.56	554	0.00	0.876	0.104

Note. ATT = attitudes; SN = subjective norms; SE = self-efficacy; CON = controllability

Table 18

Full Structural Modeling Results for Model 2

Factor	Item	Factor Loading	*SE*	*p*	Residual Variance
ATT	A1	.815		0.00	.336
	A3	.854	.041	0.00	.271
	A4	.826	.042	0.00	.318
	A5	.919	.044	0.00	.155
	A6	.753	.049	0.00	.433
	A7	.864	.039	0.00	.254
	A8	.910	.038	0.00	.172
	A9	.690	.046	0.00	.524
SN	SN1	.555		0.00	.692
	SN2	.904	.174	0.00	0.184
	SN5	.746	.132	0.00	.443
	SN6	.807	.145	0.00	.349
	SN7	.815	.149	0.00	.336
	SN8	.695	.130	0.00	.517
SE	SEC1	.676		0.00	.543
	SEC2	.792	.086	0.00	.372
	SEC3	.696	.078	0.00	.515
	SEC4	.800	.086	0.00	.359
	SEC5	.771	.088	0.00	.406
INT	I1	.833		0.00	.307
	I2	.943	.059	0.00	.111
	I5	.508	.066	0.00	.742
	I6	.921	.044	0.00	.151
	I7	.879	.044	0.00	.228

Note. ATT = attitudes; SN = subjective norms; SE = self-efficacy; INT = intentions

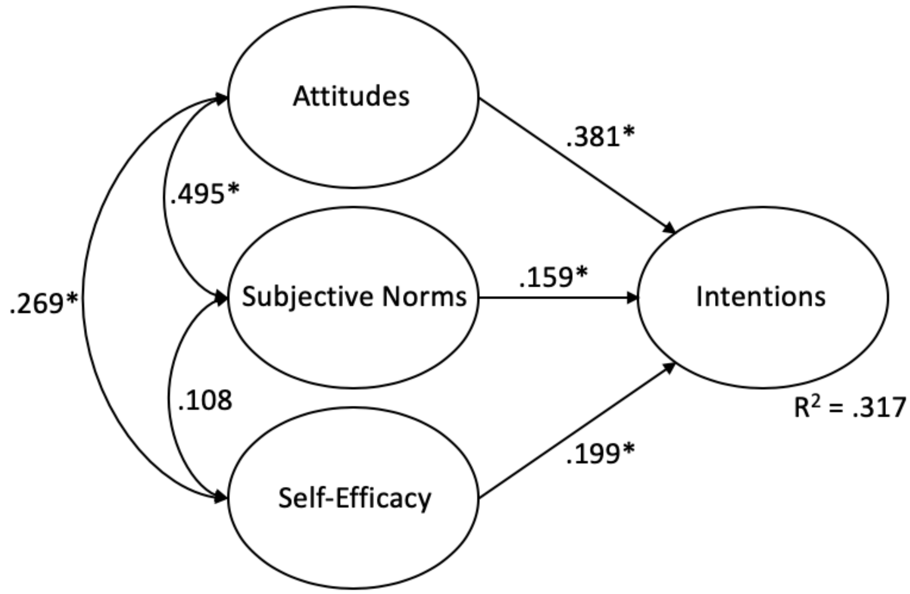

Figure 21. Structural Modeling Results for Model 2.
$*p < .05$

Chapter Five – Discussion

Legally required of all special educators (IDEA, 2004), and evidenced to be associated with more effective instructional planning (Fuchs et al., 1989; Stecker & Fuchs, 2000) and improved outcomes for students with disabilities (Fuchs et al., 1984; Mirkin et al., 1982; Stecker & Fuchs, 2000), frequent IEP data collection is critical to the provision of effective individualized services in special education. Like so many other evidence-based practices, however, a gap exists between what we know, what teachers reportedly believe, and what teachers reportedly practice in regards to data collection (Brawley & Stormont, 2013; Ruble et al., 2018; Sandall et al., 2004), leaving us to question what factors are facilitating the persistence of inconsistent and relatively subjective data collection methods. Because the decision of whether and how to engage in IEP data collection is ultimately up to each individual teacher, a focus on uncovering the impact of individual-level factors inherent in teachers is a logical first step.

Individual's beliefs and their subsequent motivation to engage in a practice has been linked to the degree to which the practice is both accepted and implemented (e.g., Aarons, Hurlburt, & McCue Horwitz, 2011; Han & Weiss, 2005; Klingner, Ahwee, Pilonieta, & Menendez, 2003; Michie, et al., 2011; Sparks 1988). More specifically, intentions—often conceptualized as a person's motivation or commitment to engage in a behavior—have been found to play a critical role in implementation (e.g., Ajzen, 1985; Han and Weiss, 2005; Michie et al., 2011; Schwarzer, 2008). Simply put, those who form an intention to engage in IEP data collection are more likely to do so. Given the promising ability of the Theory of Planned Behavior (TPB; Ajzen, 1985) in organizing research aimed to better understand and predict teachers' behaviors by measuring the belief to intention pathway, it was chosen as the framework on which teachers' intentions to engage in IEP data collection was measured in this study. Utilization of a theory to guide the exploration of teachers' beliefs and intentions is not enough, however. Greater attention needs to be paid to the way in which these hard-to-define, unobservable variables are measured in order to increase the validity of interpretations made within and across studies, thus facilitating the creation of tailored implementation supports aimed at strengthening educators' ability and commitment to engage in IEP data collection.

With a focus on enhanced measurement, a necessary first step to reducing the research-to-practice gap in relation to IEP data collection, this study sought out to evaluate the dimensionality and quality of the IEP Data Collection Intentions Scale (IDCIS). The three main goals of this study were to determine 1) the extent to which the IDCIS's subscales represented four separate constructs (attitudes, subjective norms, perceived behavioral control, and intentions); 2) which items on the IDCIS serve as appropriate indicators of each construct, such that all levels of intentions can be validly and reliably measured; and 3) whether the TPB serves as an appropriate theoretical model to measuring teacher's IEP data collection intentions. Following a discussion of all key findings, study limitations and suggested directions for future research will be highlighted.

Dimensionality and Quality of the IDCIS Subscales and Corresponding Indicators

In analyzing the dimensionality of the IDCIS, results of the single-factor CFAs and subsequent multi-factor CFA suggests that the scale represents five separate constructs including teachers' attitudes, subjective norms, self-efficacy, controllability, and intentions related to IEP data collection. Furthermore, item analysis and Rasch modeling procedures confirmed the quality of each set of indicators encompassing the five IDCIS subscales, suggesting that valid and reliable interpretations about teachers' IEP data collection intentions can be based on most IDCIS scores. In summary, item analyses resulted in high positive item-total correlations (.65 and above), meaning teachers responses were typically correlated with their overall score and items did a good job of discriminating between teachers who possess similar levels of each construct. Moreover, all items with the exception of A9 had infit and outfit values that fell within the range for productive measurement as defined by Linacre (2012), suggesting the IDCIS items serve as useful indicators in the measurement of each construct. Item characteristic curves (ICCs) for most items highlighted the proper scaling of response options, while the item thresholds and Wright maps demonstrated each subscale's ability to tap all levels of each measured construct.

Following is a discussion of the results as they relate to the broader behavior change literature as well as to the previous measurement of constructs known to impact teachers' data collection practices. Given that the outcome of these analyses was predominantly positive and straightforward, the discussion will focus primarily on unexpected and relatively poor results, highlighting reasoning for subsequent decision making and specific recommendations for future scale improvement.

Attitudes. Teachers are more motivated to adopt a practice if they find it important (Sparks, 1988) and believe it will benefit their students (Klingner et al., 2003). A general belief in the importance of data collection, however, is not sufficient in facilitating the frequent and systematic collection of data (Brawley & Stormont, 2013; Ruble, McGrew, Wong, & Missall, 2018; Sandall et al., 2004). Therefore, in measuring teachers' attitudes (as a representation of their behavioral beliefs) toward IEP data collection, the IDCIS's Attitudes subscale included questions related to perceptions of the importance and utility of data collection as well as the feelings one experiences when thinking about data collection.

Though the CFA of all nine original items (A1-A9) resulted in fit statistics suggesting good fit, item A2 (*How useful, if at all, is IEP data collection?*) was removed due to redundancy; the specific utility of data collection was targeted in items A3-A8. The final Attitudes subscale captured teachers' perceptions of the importance and utility of IEP data collection, as well as the feelings one experiences when thinking about IEP data collection. While this was a slight deviation from Ruble and colleagues (2018) conceptualization, as their measurement of attitudes only included items related to importance and utility, it was in alignment with Ajzen's description of attitudes—"the degree to which a person has a favorable or unfavorable evaluation or appraisal of the behavior" (1991, p.188). Additionally, inclusion of this item was also supported by Francis and colleagues' (2004) recommendation to include an experiential item (i.e., a question aimed to capture how it feels to perform the behavior) in the measurement of attitudes.

As described in the results, item A9 (*To what extent, if at all, do you agree that thinking about IEP data collection gives you a positive feeling?*) was retained irrespective of its slightly inflated infit and outfit values because of its location on the Attitudes continuum; it was created to target those with the highest level of attitudes toward IEP data collection. Moreover, after further review of this item, its adherence to writing guidelines was questioned, as it was possibly biased in its sole inclusion of the word "agree" in the item stem, which represents only one end of the bipolar scale represented in the response options. Though it is unclear if this oversight impacted participants' responses to item A9, future modifications to the IDCIS should include the revision of this item.

Subjective Norms. In addition to a leader's belief and commitment to any given practice, an individual who is dedicated to a practice and promotes the efficacy of the practice among their colleagues is thought to promote implementation (e.g., Taylor et al., 1999; Fixsen, et al., 2005). Because a person's behavior—what they talk about, attend to, and reward—can serve as a proxy for their beliefs, thus impacting the organizational climate and the behaviors of others (Schein, 2004), the Subjective Norms subscale was originally designed to measure teachers' observations of others' behaviors that might be better representations of their actual beliefs. This was thought to be an improvement to Ruble and colleagues' (2018) conceptualization of subjective norms, as they focused solely on teachers' perceptions of others' beliefs of the importance of data collection.

Results of the CFA suggested the item related to parents' perceptions of IEP data collection (item SN3) did not contribute to the measurement of the construct, therefore, the final subscale only captured teachers' perceptions of the importance of IEP data collection only through the eyes of their administrators and fellow service providers. Necessary removal of SN3 was not surprising, however, as it was hypothesized that the data collection beliefs of parents would be impacted by teachers' data collection practices. For example, parents who were accustomed to seeing and hearing about data might exhibit more behaviors suggesting a perception of importance (or might have more opportunity to exhibit behaviors associated with any degree of importance), while parents who rarely see or hear about data might exhibit more behaviors suggesting a perception of little to no importance (or might not ever have an opportunity to behave in a way that would be linked to the importance of data collection). Because parents' perception of the importance of IEP data collection likely impacts the "social pressure" teachers' experience associated with data collection—an important aspect of the construct as defined by Ajzen (1991)—teachers' perceptions of parents' beliefs should be included. Future modifications to this subscale should include an item related to teachers' perceptions of parents' beliefs that is not contingent on the teacher's data collection behaviors.

Additionally, the item created to gather teachers' observations of their coworker's IEP data collection practices (item S4: *How often, if ever, do you observe your coworkers [i.e., other ECSE teachers and related service providers] engaging in IEP data collection?*) was removed due to its relatively low factor loading. Because an individual's overt behaviors are known to impact organizational climate regardless of the unspoken rules and/or values that exist (i.e., Schein, 2004), this item was designed to serve as an additional indicator of the views held by teachers' coworkers. Furthermore, it was viewed as an important indicator of subjective norms, as it was hypothesized to provide information to corroborate the data from item S2 (*How important, if at all, is IEP data collection to your coworkers [i.e., other ECSE teachers and related service providers?])*. After further review of the language used in item S4, however, multiple interpretations were uncovered. While the item was written to capture what teachers "see" other services providers doing related to data collection while "naturally" in the presence of these people, this item could have been interpreted as the frequency in which teachers engage in more structured observations of their coworkers with the sole purpose of learning more about their data collection practices. Additionally, item A4 was not sensitive to the various degrees of interactions with other service providers experienced by ECSE teachers. While many teachers spend their days surrounded by other service providers, others might go weeks before spending time in a classroom with another service provider. Given the hypothesized importance of this item, future modifications of the IDCIS should include an indicator related to what data collection behaviors teachers "see" other service providers engaging in that is applicable to all teachers, regardless of the rate at which they in the presence of other service providers.

Another necessary modification was highlighted by the ICCs for items SN5-SN7. As displayed in Figure 5, the rating scales for these items did not function well; the probability of choosing "yearly" was only higher than the probability of choosing one of the other three response options for a very small range of theta. To improve the functioning of the response options, it is recommended that the response options for these items be modified. Instead of "never", "yearly", "quarterly", and "monthly" as included in the current version of the IDCIS, never and yearly should be combined and all other options should be equally spaced from yearly to monthly.

Finally, the removal of items SN3 (*How important, if at all, is IEP data collection to your students' parents?*) and SN4 (*How often, if ever, do you observe your coworkers engaging in IEP data collection?*) likely impacted this subscale's person reliability and person separation values (0.80 and 1.99, respectively), as more items that are targeted to the level of theta represented in the sample are reported to increase these values (Linacre, 2012). The future replacement of these items with two high quality items (based on the previous discussion of this subscale's CFA results) is recommended to improve the scale's overall functioning.

Perceived Behavioral Control. As highlighted by Armitage and Conner's (2001) review of 185 studies theoretically rooted in the TPB, discrepancies exist between how perceived behavioral control has been defined, with some researchers focusing more on internal confidence, others focusing more on external barriers, and yet another group focusing more on perceived difficulty. Based on this finding and aligned with Ajzen's (1991) definition—the perceived ease or difficulty associated with engaging in the behavior, which is thought to be impacted by past experiences in addition to anticipated barriers—the initial measure of perceived behavioral control was multifaceted. Following scale development, the IDCIS included 10 items aligned with self-efficacy (i.e. confidence and perceived difficulty), 17 items aligned with controllability (i.e. the extent to which specific barriers impact data collection practices), and 3 items related to general perceptions of behavioral control. Given the sheer number of items as well as their alignment with the many interpretations of perceived behavioral control, the necessary division of these items into more than one construct based on the CFA results was not surprising.

Self-Efficacy. Consistent with Bandura's notion of self-efficacy (1971), a teacher might have a positive attitude toward IEP data collection—including the belief that data collection is both important and useful—but if they do not perceive themselves as possessing the capacity necessary to effectively engage in the frequent and systematic collection of IEP data, the teacher is hypothesized to put forth less effort and identify more barriers. Teachers who are confident in their ability to collect IEP data, however, are thought to be better equipped to overcome barriers, making future data collection less effortful. Although Francis and colleagues' (2004) defined self-efficacy as a combination of confidence and perceived difficulty in their manual for the construction of surveys grounded in the TPB, all five items related to perceived difficulty were dropped given their relatively limited ability to provide novel information; individuals' perception of difficulty was thought to be captured in the items related to the impact of common barriers. Furthermore, while the overall fit of the model including the five items related to confidence (i.e., items SEC1-SEC5) was less than adequate when the conventional cut-off values were used, these items were retained due to their alignment with Bandura's (1986) conceptualization of self-efficacy. As opposed to general confidence related to a domain, Bandura reported that behavior is better predicted by one's beliefs about their capacity to engage in all behaviors necessary for success (Bandura, 1986; Bandura, 2006); which in this case, is everything from writing clear and measurable IEP objectives to collecting IEP data on a daily basis. This was viewed as an improvement to Ruble and colleague's (2018) two items related to individual's confidence in collecting and using data, as the collection and use of data are comprised of numerous discrete behaviors.

Controllability. Based on previous reports, the main contextual factor impacting the data collection practices of ECSE professionals is lack of time (Banjeree & Luckner, 2013; Cooke et al., 1991; Sandall et al., 2004). The range in educational settings that exist in ECSE and the even wider range of instructional contexts (curriculum, nature of classroom routines, teacher roles and responsibilities, adult-to-child ratio, the nature of children's delays/disabilities, etc.) in which teachers are expected to operate within present additional barriers potentially unique to IEP data collection in ECSE. With the goal of facilitating the development of individualized implementation supports that enable teachers to persist and succeed with their data collection efforts, the IDCIS originally included 20 items related to controllability.

Immediate removal of the three items originally designed as indicators of one's general perceptions of the control they have over their own data collection behaviors (PBC1-PBC3) was supported not only by their low factor loadings, but also by Armitage and Conner's (2001) findings that general perceptions of control over behavior were associated with significantly weaker correlations to both intentions and behavior when compared to items that measured confidence or difficulty. Of the remaining 17 items (B1-B17), B3 (*How often, if ever, does the availability of time to analyze and interpret IEP data decrease your ability to engage in daily IEP data collection?*) was dropped as no other items on the IDCIS were related to teachers' use of data. After the removal of these four items, the final Controllability subscale resulted in less than adequate fit according to the criteria most commonly used. While better fit could have been achieved with less items, the decision to retain all 16 items was supported by positive item analysis and item response modeling results as well as the intended uses of the IDCIS. While additional attention toward the measurement of this construct is warranted, the Controllability subscale was viewed as an improvement to the scale used by Ruble et al. (2018), which focused solely on the degree to which lack of time, unclear measurement systems, and too many students impacted teachers' data collection practices.

Intentions. In applying the TPB (Ajzen, 1985) to this context, teachers' intentions to engage in IEP data collection are thought to represent the motivational factors influencing their future data collection behaviors; those with stronger data collection intentions are thought to put forth more effort when it comes to collecting data than those with weak intentions. Improving upon the one-item measure of intentions (i.e., *I intend to keep data over the next 2 weeks*) used by Ruble et al. (2018), the original Intentions subscale included in the IDCIS was comprised of seven items, with items specifically written to target a wide range of intentions from teachers who primarily intend to rely on memories of student performance to those who primarily intend to engage in systematic documentation of student performance based on direct observation.

Though the CFA results supported the removal of items I3 (i.e., the extent to which one intends to reflect on their memory of students' skills while writing down informal notes) and I4 (i.e., the extent to which one intends to talk with others about their observations of students' skills while writing taking informal notes), additional evidence supporting their removal was desired because of both items were hypothesized to be important contributors to the construct—these items were written to tap the lowest levels of intentions. After sequential review of all seven items—from low levels of intentions to high levels of intentions—it was determined that while items I3 and I4 represented behaviors that teachers low on the intentions continuum could easily endorse, teachers with high levels of intentions (those who intend to engage in more objective methods represented in items I6 and I7) would likely struggle to endorse these highly subjective methods. Because scale development was rooted in construct modeling, this was viewed as a major flaw and items I3 and I4 were removed.

The final five items chosen to represent the Intentions subscale of the IDCIS (i.e., I1, I2, I5-I7) did not result in adequate fit based on the cut-off criteria traditionally utilized in CFA research. Removal of item I2 (i.e., *To what extent, if at all, is IEP data* *collection a high priority for you over the next month?)* significantly improved the model and resulted in overall measures of fit that exceeded all conventional cut-off values, this item was retained given the impact of goal priorities on the strength of the intention-behavior relationship (Conner et al., 2016). Moreover, the retention of this item was supported by the results of all subsequent analyses.

Evaluating the Structural Model

There are many factors—both internal and external—that are thought to impact ECSE teachers' IEP data collection practices. Given that the choice of whether and how to engage in IEP data collection is ultimately left up to each individual teacher, the examination of teachers' intentions, an important individual-level factor known to impact implementation (e.g., Armitage & Conner, 2001; Han & Weiss, 2005; Michie et al., 2011; Schwarzer, 2008) is a logical first step. The TPB (Ajzen, 1985) serves as a well-organized and evidence-based framework for exploring intentions and their influence on future behavior, while at the same time acknowledging and accounting for the many external factors hypothesized to impact one's commitment to engage in a future behavior. As such, the appropriateness of the TPB in grounding the measurement of ECSE teachers' IEP data collection intentions was explored.

Based on the results of the structural modeling, Model 2, which included self-efficacy over controllability, resulted in the best fit. In this model, 32% of the variance in intentions was accounted for by attitudes, subjective norms, and self-efficacy. This is consistent with Armitage and Conner's (2001) findings that attitudes, subjective norms, and perceived behavioral control accounted for an average of 39% of the variance in behavioral intentions across 161 studies. Furthermore, this is comparable to Ruble and colleagues' (2018) findings that 30% of the variance in intentions to collect IEP data could be predicted using their measures of attitudes, subjective norms, self-efficacy, perceived behavioral control, and administrative support. Considering the results of the CFAs and Rasch modeling that revealed the need for future modifications to several IDCIS subscales—modifications that are hypothesized to provide additional validity evidence based on the scale content (i.e., construct representation and content coverage)—these results are thought to represent a low estimate of the TPB's ability to facilitate our understanding of teachers' intentions to engage in future IEP data collection.

Regardless of model fit, the exclusion of the Controllability subscale requires future attention. Though the measurement of perceived behavioral control is commonly replaced with a measure of self-efficacy—a substitution supported by Ajzen's (1991) description of the compatibility of these two constructs—the IDCIS's Self-Efficacy subscale only includes items serving as indicators of confidence, which is hypothesized to have a detrimental impact on the scale's ability to inform the creation of future implementation supports. For example, knowing that a teacher is only slightly confident in their ability to collect IEP data on a daily basis does not lead to a better understanding of what specific skill(s) (planning for IEP data collection, time management, accessing or modifying existing tools, writing IEP objectives, collecting data across all developmental domains, individualizing data collection methods to meet students' needs, etc.) should be targeted to increase their confidence. As the IDCIS is modified based on the results of this study, future explorations of the usefulness of the Controllability subscale in explaining intentions and predicting subsequent data collection behavior is warranted.

Limitations and Future Directions

As previously described, claims underlying the IDCIS's Interpretation and Use Argument (IUA) include both those about the constructs being measured as well as how the IDCIS scores will be interpreted and used. Based on these claims and designed to support the sequence of inferences necessary to move from a teacher's response to each individual item to the final conclusions drawn about that teacher, a specific set of procedures were followed to generate necessary validity evidence. As such, it is only appropriate to base the discussion of study limitations on the extent to which each limitation impacted the evidence supporting this sequence of inferences. Limitations previously highlighted in the results chapter will be briefly reviewed, while other limitations will be discussed more thoroughly. Regardless of the extent to the discussion, each limitation will be linked to recommendations for future research in order to accelerate IEP data collection practices that are more closely aligned with evidence-based practices and teachers' reported beliefs.

Limitations rooted in scale development. Measurement error, a common source of survey error produced by inaccurate responses irrespective of intentionality, was identified as the main study limitation; one that was produced by the decision-making process that facilitated the IDCIS development. Regardless of the effort made to ensure item quality, including the utilization of the TPB (Ajzen, 1985) to establish clearly defined constructs as well as the careful application of evidence-based item-writing guidelines (Dillman et al., 2014), limitations rooted in item construction that were thought to impact the scoring and extrapolation inferences exist. As such, the future measurement of teachers' attitudes toward IEP data collection should be cognizant of possible bias in items designed to capture teachers' feelings toward data collection. Additionally, the measurement of teachers' subjective norms related to IEP data collection should include items designed to gather teachers' perceptions of how important data collection is to their students' parents that are less dependent on their own data collection practices as well as clearly written items designed to gather teachers' observations of their coworker's data collection practices that are sensitive to the varying frequencies in which ECSE teachers work with students alongside other service providers.

Furthermore, given that scale development was grounded in construct modeling, limitations rooted in item design that were uncovered by the Rasch analyses potentially impacted the scoring and extrapolation inferences. Future modifications should include additional items designed to tap those with highest levels of attitudes and self-efficacy. In addition, the measurement of teachers' observations of the behaviors of ECSE leaders should include items with response options that are scaled to better represent the frequency in which teachers are in the presence of such leaders. Finally, subsequent measurement of ECSE teachers' data collection intentions should ensure proper application of a construct map, especially when designing items to target those with the weakest intentions.

By attending to all aforementioned recommendations, the measurement of each construct represented in the IDCIS will include a greater number of indicators that are tailored to the range of each trait possessed by ECSE teachers. In addition to supporting the scoring and extrapolation inferences, this inclusion of more items individualized to the sample will likely improve person reliability and separation values, both of which were lower than preferred for the Subjective Norms and Intentions subscales. As a result, additional validity evidence supporting the assumption that a teacher's score represents their average score based on repeated measurement—the generalization inference—will be generated.

Limitations rooted in the Theory of Planned Behavior. While the TPB (Ajzen, 1985) was instrumental in the creation of the IDCIS in addition to supporting subsequent production of validity evidence supporting the interpretation and use of IDCIS scores, several limitations have been identified that are thought to stem from the study's utilization of the theory. First, as directly stated in Armitage and Conner's (2001) review and implicitly highlighted in the discrepancies that exist between more recent applications of the theory, questions remain regarding how each construct is defined and how variations in definitions impact the relations between constructs. Specific to this study, less than adequate model fit for indicators of self-efficacy and controllability is viewed as a limitation likely impacting the extrapolation inference. Specific attention to how these constructs are defined and measured is warranted in order to support future explorations of teachers' intentions, both those specific to and independent of IEP data collection.

While the sole measurement of constructs included in the TPB was viewed as a feasible first step in the creation of a scale intended to measure and explain ECSE teachers' data collection intentions, thus informing the creation of effective implementation supports, there are additional variables known to impact one's intentions that were not addressed in the current study. Because the creation and validation of the IDCIS was grounded in specific claims about the constructs being measured, this limitation is not thought to impact the series of inferences necessary for the valid interpretation and use of IDCIS scores, but it is viewed as a major restriction in the utility of the IDCIS in reducing the research-to-practice gap. To address this limitation, IDCIS modifications and recommended uses will be discussed here.

First, as recognized in the TPB literature, an individual's reported self-efficacy needs to be a realistic reflection of their genuine capabilities in order to aid in the accurate explanation and/or prediction their behaviors (Ajzen, 1991). A future modification to the IDCIS, therefore, would be the addition of a subscale designed to measure actual knowledge and skills related to IEP data collection in hopes to better understand teachers' attitudes and perception of controllability related to data collection. Second, the TPB does not account for the role that one's past behavior plays in the formation of intentions. Given the hypothesized impact of habitual behavior on the constructs measured in the TPB (Ouellette & Wood, 1998), the collection of objective data illustrating teachers' behaviors related to IEP data collection preceding IDCIS use is recommended. While five demographic questions created to gather teachers' self-reports of data collection practices over the past two weeks were included in the survey administered in this study, these items mirrored the indicators of intentions—specifically targeting the frequency in which teachers rely on their memories, conversations with others, anecdotal notes taken during informal observations, and relatively more objective data collected during structured observations—and therefore experienced similar problems during all attempted CFA and Rasch analysis procedures. Because of this, and recognizing the potential for self-reported practices to be an inaccurate representation of actual practices, the impact of past data collection behavior on teachers' attitudes, self-efficacy, controllability, and intentions were not explored. Future attention toward the measurement of these variables will help in the identification of and differentiation between skill and performance deficits, further aiding in the creation of tailored implementation supports.

A final limitation in the utilization of the TPB in exploring ECSE teachers' data collection intentions is the theory's failure to account for an individual's satisfaction with their current practices, thus uncovering potential demand or readiness for change. According to the Transtheoretical Model (TTM; Porchaska, DiClemente, & Norcross, 1993), successful implementation of any practice requires individuals to progress through a variety of stages. For example, as described by Porchaska et al. (1993), individuals in the *precontemplation* stage have no intention to change their behavior in the near future, while those in the *contemplation* stage recognize the need for change but have yet to commit to it. Furthermore, those in the *preparation* stage have initiated their commitment to change and have already started making steps toward the implementation of a new practice. In addition to the objective measurement of past data collection behavior, measuring one's satisfaction with their current data collection practices will help in determining whether the utility of the TPB is equivalent across all stages of change, or whether it is more useful in explaining the intentions and predicting the behaviors of teachers in one particular stage over another.

Conclusion

In closing, following minor adjustments previously discussed, the IEP Data Collection Intention Scale (IDCIS) can be used to produce precise measures of teachers' attitudes, subjective norms, self-efficacy, controllability, and intentions related to the collection of IEP data. Furthermore, the scores produced by IDCIS administration can be used to make valid and reliable inferences about teachers' levels of each construct in order to inform the creation and modification of future implementation supports. While the TPB (Ajzen, 1985) serves as an appropriate theoretical framework guiding initial explorations of ECSE teachers' IEP data collection intentions, additional attention is warranted to determine if and how the Controllability subscale can be used to better understand teachers' intentions. Furthermore, when used concurrently with a measure of teacher's current data collection behaviors, their satisfaction with their current practices, and their actual knowledge related to various evidence-based data collection practices, the IDCIS will be instrumental in the prediction and explanation of IEP data collection practices following the application of various implementation supports.

References

Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The evidence-based practice attitude scale (EBPAS). *Mental Health Services Research*, *6*(2), 61–74.

Akers, L., Del Grosso, P., Atkins-Burnett, S., Boller, K., Carta, J., & Wasik, B. A., (2014). Tailored teaching: Teachers' use of ongoing child assessment to individualize instruction (Volume II). Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, 2014.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action Control* (pp. 11–39). Berlin, Heidelberg: Springer Berlin Heidelberg.

Armitage, C. J., & Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of social and clinical psychology*, *4*(3), 359-373.

Bandura, A. (1971). *A social learning theory*. Morristown, N.J.: General Learning Press.

Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-efficacy beliefs of adolescents*, *5*(1), 307-337.

Banerjee, R. & Luckner, J. (2013). Assessment practices and training needs of early childhood professionals. *Journal of Early Childhood Teacher Education, 34*(3), 231-248.

Beaujean, A. A. (2012). BaylorEdPsych: R package for Baylor University educational psychology quantitative courses. R package version 0.5, URL http://CRAN.R-project.org/package=BaylorEdPsych.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. doi:10.1037/0033-2909.107.2.238

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*(1), 78-117.

Boardman, A. G., Argüelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers' views of research-based practices. *The Journal of Special Education*, *39*(3), 168–180.

Brawley, S., & Stormont, M. A. (2014). Investigating reported data practices in early childhood: An exploratory study. *Journal of Positive Behavior Interventions*, *16*(2), 102-111.

Bricker, D., Cripe, J., & Slentz, K. (2003). *Assessment, evaluation, and programming system (AEPS).* Baltimore, MD: Paul Brookes Publishing.

Buzhardt, J., Walker, D., Greenwood, C. R., & Heitzman-Powell, L. (2012). Using technology to support progress monitoring and data-based intervention decision making in early childhood: Is there an app for that? *Focus on Exceptional Children*, *44*(8), 1-18.

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. Journal of Statistical Software, 48(6), 1-29. Retrieved from doi:10.18637/jss.v048.i06.

Christle, C. A., & Yell, M. L. (2010). Individualized education programs: Legal requirements and research findings. *Exceptionality*, *18*(3), 109–123.

Conner, M., Abraham, C., Prestwich, A., Hutter, R., Hallam, J., Sykes-Muskett, B., Morris, B. & Hurling, R. (2016). Impact of goal priority and goal conflict on the intention–health-behavior relationship: Tests on physical activity and other health behaviors. *Health Psychology*, *35*(9), 1017-1026.

Cooke, N. L., Heward, W. L., Test, D. W., Spooner, F., & Courson, F. H. (1991). Student performance data in the classroom measurement and evaluation of student progress. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *14*(3), 155-161.

Cook, B. G., & Cook, S. C. (2013). Unraveling evidence-based practices in special education. *The Journal of Special Education*, *47*(2), 71–82.

David, D., Miclea, M., & Opre, A. (2004). The information-processing approach to the human mind: Basics and beyond. *Journal of Clinical Psychology, 60* (4), 353-356.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219–232.

Deno, S. L., & Mirkin, P. K. (1977). Data-based program modification: A manual. Council for Exceptional Children (Publication Sales Unit), Reston, VA.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Division for Early Childhood (DEC). (2007). *Promoting positive outcomes for children with disabilities: Recommendations for curriculum, assessment, and program* *evaluation* (Position paper). Missoula, MT: Division for Early Childhood, Council for Exceptional Children.

Division for Early Childhood. (2014). DEC recommended practices in early intervention/early childhood special education 2014. Retrieved from http://www.dec-sped.org/recommendedpractices.

Dowd, E. T., Clen, S. L., & Arnold, K. D. (2010). The specialty practice of cognitive and behavioral psychology. *Professional Psychology: Research and Practice, 41* (1), 89-95.

Downs, A., & Strand, P. S. (2006). Using assessment to improve the effectiveness of early childhood education. *Journal of Child and Family Studies*, *15*(6), 671-680.

Doyle, W., & Ponder, G. A. (1977). The practicality ethic in teacher decision-making. *Interchange*, *8*(3), 1–12.

Dudley-Marling, C. (1985). Perceptions of the usefulness of IEP by teachers of learning disabled and emotionally disturbed children. *Psychology in the Schools, 22*, 65-67.

Education for All Handicapped Children Act. Public Law No. 94-142 (1975).

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M. & Wallace, F. (2005). Implementation research: A synthesis of the literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Foster, R. (2014). Barriers and Enablers to Evidence-Based Practices. *Kairaranga*, *15*(1), 50-58.

Fox, L., Veguilla, M., & Perez Binder, D. (2014). *Data decision-making and program-wide implementation of the pyramid model* (Roadmap to Effective Intervention

Practices, No. 7). Tampa: University of South Florida Technical Assistance Center on Social Emotional Intervention for Young Children.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, *57*(6), 488–500.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, *21*(2), 449-460.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, *41*(1), 93–99.

Fuchs, L. S., Fuchs, D., & Stecker, P. M. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities*, *22*(1), 51-59.

Fuchs, L. S., Fuchs, D., & Warren, L. M. (1982). Special education practice in evaluating student progress toward goals (Research Report No. 81). Minneapolis: Institute for Research on Learning Disabilities.

Garrett, J. N., & Kelley, M. F. (2000). Early childhood special education: Workplace realities. *Childhood Education*, *76*(5), 267–276.

Gersten, R., & Smith-Jones, J. (2001). Reflections on the research to practice gap. *Teacher Education and Special Education*, *24*(4), 356-61.

Gettinger, M., & Stoiber, K. (2007). Applying a response-to-intervention model for early literacy development in low-income children. *Topics in Early Childhood Special Education, 27,* 198–213.

Greenwood, C.R., Carta, J.J., Goldstein, H., Kaminski, R.A., McConnell, S.R. & Atwater, J. (2015). The Center for Response to Intervention in Early Childhood: Developing evidence-based tools for a multi-tier approach to preschool language and early literacy instruction. Journal of Early Intervention, 36, 246-262.

Greenwood, C. R., Carta, J. J., & McConnell, S. (2011). Advances in measurement for universal screening and individual progress monitoring of young children. *Journal of Early Intervention*, *33*(4), 254-267.

Gregg, D. (2009). Developing a collective intelligence application for special education. *Decision Support Systems, 47,* 455—465.

Hornby, G., Gable, R. A., & Evans, W. (2013). Implementing evidence-based practice in education: What international literature reviews tell us and what they don't. *Preventing School Failure: Alternative Education for Children and Youth*, *57*(3), 119–123.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Jones, E. D., & Krouse, J. R. (1988). The effectiveness of data-based instruction by student teachers in classrooms for pupils with mild learning handicaps. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *11*(1), 9–19.

Kasprzak, C., Hurth, J., Rooney, R., Goode, S. E., Danaher, J. C., Whaley, K. T., Ringwalt, S., & Cate, D. (2012). States' accountability and progress in serving young children with disabilities. *Topics in Early Childhood Special Education*, *32*(3), 151–163.

Klingner, J., Ahwee, S., Pilonieta, P., & Menendez, R. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children, 69*(4), 411-429.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. Journal of the American Statistical Association, 83(404), 1198--1202.

Margolis, H., & Truesdell, L. A. (1987). Do special education teachers use IEPs to guide instruction? *The Urban Review*, *19*(3), 151–159.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research.* RAND Corporation. Retrieved from https://www.rand.org/pubs/occasional_papers/OP170.html.

Marsh, H.W., Hau, K. & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11:3, 320-341, DOI: 10.1207/s15328007sem1103_2

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149-174.

McConnell, S. R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into our future. *Topics in Early Childhood Special Education*, *20*(1), 43-48.

McConnell, S. R., Wackerle-Hollman, A., Bradfield, T. A., & Rodriguez, M. C. (2012). *Individual Growth and Development Indicators of Early Literacy* (2nd ed.). St Paul MN: Early Learning Labs, Inc.

McConnell, S.R., Wackerle-Hollman, A.K., Roloff, T.A., & Rodriguez, M. (2015). Designing a measurement framework for response to intervention in early childhood programs. Journal of Early Intervention, 36, 263-280.

Meisels, S. J., Jablon, J., Marsden, D. B., Dichtelmiller, M. L., & Dorfman, A. (2001). The Work Sampling System (4th ed.). Ann Arbor, MI: Rebus Inc.

Meyer, J. (2014). *Applied measurement with jMetrik*. New York, NY: Routledge.

Minnesota Department of Education, (2018). Retrieved from http://w20.education.state.mn.us/MDEAnalytics/DataTopic.jsp?TOPICID=455.

Minnesota Professional Educator Licensing and Standards Board, (2018). Retrieved from http://w20.education.state.mn.us/MDEAnalytics/PELSB.jsp?TOPICID=437).

Minnesota Rule, 8710.5500, 2017, retrieved from: https://www.revisor.mn.gov/rules/pdf/8710.5500/2017-08-21%2012:12:30+00:00.

Mirkin, P., Deno, S., Tindal, G., & Kuehnle, K. (1982). Frequency of measurement and data utilization as factors in standardized behavioral assessment of academic skill. *Journal of Behavioral Assessment*, *4*(4), 361-370.

Morgan, D. P., & Rhode, G. (1983). Teachers' attitudes toward IEP's: A two-year follow-up. *Exceptional Children*, *50*(1), 64–67.

National Center on Intensive Intervention. (2013). *Data-based individualization: A framework for intensive intervention.* Washington, DC: Office of Special Education, U.S. Department of Education.

Notari, A. R., & Bricker, D. D. (1990). The utility of a curriculum-based assessment instrument in the development of individualized education plans for infants and young children. *Journal of Early Intervention*, *14*(2), 117-132.

Odom, S. L. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*(1), 53-61.

Odom, S. L., & Wolery, M. (2003). A unified theory of practice in early intervention/early childhood special education. *The Journal of Special Education*, *37*(3), 164–173.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, 124(1), 54.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307–332.

Pierce, W. D. & Cheney, C. D. (2013). *Behavior analysis and learning*. (5^{th} Ed). Psychology Press.

Qualtrics (2019). Qualtrics Survey Software. Provo, UT. Retrieved from http://www.qualtrics.com

Reeve, B. B. & Masse, L. C. (2004). Cognitive interviewing revisited: A useful technique, in theory? In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 23-44). Hoboken, New Jersey: John Wiley & Sons, Inc.

RStudio Team (2018). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/.

Ruble, L. A., McGrew, J. H., Wong, W. H., & Missall, K. N. (2018). Special education teachers' perceptions and intentions toward data collection. *Journal of Early Intervention*, *40*(2), 177–191.

Ruble, L. A., Toland, M. D., Birdwhistell, J. L., McGrew, J. H., & Usher, E. L. (2013). Preliminary study of the autism self-efficacy scale for teachers (ASSET). *Research in autism spectrum disorders*, *7*(9), 1151-1159.

Sandall, S. R., Schwartz, I. S., & Lacroix, B. (2004). Interventionists' perspectives about data collection in integrated early childhood classrooms. *Journal of Early Intervention*, *26*(3), 161-174.

Smith, S. W. (1990). Individualized education programs (IEPs) in special education—From intent to acquiescence. *Exceptional Children*, *57*(1), 6–14.

Sparks, G.M. (1988). Teachers' attitudes toward change and subsequent improvements in classroom teaching. *Journal of Educational Psychology*, *80*(1), 111-117.

Stecker, P. M., & Fuchs, L. S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research & Practice*, *15*(3), 128-134.

Suarez-Balcazar et al., 2010

Sugai, G., & Horner, R. R. (2000). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review*, *35*(2), 245–259.

Schwarzer, R. (2008). Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. *Applied Psychology: An International Review*, 57, 1–29.

Taylor, L., Nelson, P., & Adelman, H. S. (1999). Scaling-up reforms across a school district. *Reading and Writing Quarterly, 15*, 303-325.

Teaching Strategies, Inc. (2012). *Teaching Strategies GOLD birth through kindergarten touring guide.* Washington, DC: Author.

Thorndike, R. M., & Thorndike-Christ, T. M. (2010). *Measurement and evaluation in psychology and education, 8th Edition.* Boston, MA: Pearson Education, Inc..

U.S. Department of Education (2015). 36th Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act, 2014. Retrieved from http://www2.ed.gov/about/reports/annual/osep/2014/parts-b-c/36th-idea-arc.pdf.

Vort Corporation. *Hawaii Early Learning Profile® (HELP® 3–6),* Second edition. Palo Alto, CA: Vort Corporation, 2010.

Wackerle-Hollman, A., Duran, L., & Rodriguez, M. (2015). *Individual Growth and Development Indicators-Español*. St Paul, MN: Early Learning Labs.

Walker, D., Carta, J. J., Greenwood, C. R., & Buzhardt, J. F. (2008). The use of individual growth and developmental indicators for progress monitoring and intervention decision making in early education. *Exceptionality, 16* (1), 33-47.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*, 158-177.

Wesson, C. L., King, R. P., & Deno, S. L. (1984). Direct and frequent measurement of student performance: If it's good for us, why don't we do it? *Learning Disability Quarterly*, *7*(1), 45-48.

Willis, G. B. (2004). Cognitive interviewing: A tool for improving questionnaire design. Sage Publications, Inc.

Yan, Z., & Sin, K. F. (2014). Inclusive education: Teachers' intentions and behaviour analyzed from the viewpoint of the Theory of Planned Behaviour. *International Journal of Inclusive Education, 18*, 72–85.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. Retrieved from http://www.jstatsoft.org/v48/i02/.

Zweig, J., Irwin, C. W., Kook, J. F., & Cox, J. (2015). Data collection and use in early childhood education programs: Evidence from the Northeast Region (REL 2015–084). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Appendix A

Study Survey

The purpose of this study is to better understand ECSE teachers' beliefs and intentions related to IEP data collection in order to support the creation of implementation supports individualized to the barriers teachers face across the state. This survey should take approximately 15 minutes to complete. At the end of the survey, you will be given the opportunity to share any comments you may have related to the survey content.

All survey responses will be gathered anonymously. Those who complete the survey, however, will be given an opportunity to submit an email address for a chance to win one of 10 $25 Visa gift cards. If you choose to submit your email address, please know that it will be immediately deleted from the dataset containing survey responses and will be kept in a separate file, such that it will not be possible to link your responses to any identifying information.

Your participation in this study is voluntary. If you decide to participate, you may withdraw at any time. If you decide not to participate or if you choose to withdrawal, it will not affect the relationships you may have with the University of Minnesota.

If you have any questions about this study, please contact Brenna Rudolph at nolan225@umn.edu. If you would prefer to talk to someone other than the researcher, you are encouraged the contact the Research Participants' Advocate Line at 612-625-1650.

ELECTRONIC CONSENT: Clicking on the "agree" button below indicates that you have read the above information, you are at least 18 years of age, and you voluntarily agree to participate. If you do not wish to participate, please decline participation by clicking on the "disagree" button.

○ Agree (1)

○ Disagree (2)

Are you **currently employed as an ECSE teacher?**

○ Yes (1)

○ No (2)

For at least the last 2 months, have you provided services to at least one child on an IEP in a classroom setting (e.g., inclusive preschool classroom, self-contained ECSE classroom, child care classroom)?

○ Yes (1)

○ No (2)

The following 4 questions relate to your beliefs about the importance and usefulness of IEP data collection.

*For the purpose of this survey, "IEP data collection" refers to the ***ongoing collection of data highlighting student progress toward meeting IEP objectives.*** IEP data collection does NOT include the intermittent collection of data for federal, state, and/or district accountability purposes, unless directly tied to your student's IEP goals and objectives.

A1. How **important**, if at all, is IEP data collection? (*n*=368)

- ○ Very important (4) – 62%
- ○ Mostly important (3) – 32.3%
- ○ Slightly important (2) – 5.4%
- ○ Not at all important (1) – 0.3%

A2. How **useful**, if at all, is IEP data collection? (*n*=368)

- ○ Very useful (4) – 54.3%
- ○ Mostly useful (3) – 38.6%
- ○ Slightly useful (2) – 7.1%
- ○ Not at all useful (1) – 0%

To what extent, if at all, do you agree that **IEP data collection improves these things**?

	Strongly agree (4)	Agree (3)	Disagree (2)	Strongly disagree (1)
A3. Quality of my IEP objectives (*n*=368)	46.7%	47.8%	4.6%	0.8%
A4. Quality my progress reporting (*n*=368)	64.4%	33.1%	2.2%	0.3%
A5. Quality of my instruction (*n*=367)	40.3%	46.9%	11.4%	1.4%
A6. My accountability to others (*n*=367)	42.5%	51.8%	5.4%	0.3%
A7. My communication with parents (*n*=367)	45.2%	48%	5.7%	1.1%
A8. My students' outcomes (*n*=367)	43.6%	48.2%	7.1%	1.1%

A9. To what extent, if at all, do you agree that **thinking about IEP data collection gives you a positive feeling**. (*n* = 368)

- ○ Strongly Agree (4) – 16%
- ○ Agree (3) – 40.5%
- ○ Disagree (2) – 38.3%
- ○ Strongly Disagree (1) – 5.2%

The following 6 questions relate to your beliefs about the norms present within your program that impact your IEP data collection practices.

*As a reminder, "IEP data collection" refers to the ***ongoing collection*** *of data highlighting student progress toward meeting IEP objectives*.

How important, if at all, is IEP data collection to these people?

	Very important (4)	Mostly important (3)	Slightly important (2)	Not at all important (1)
SN1. ECSE leadership in your district (i.e., coordinator, program lead, and professional development coaches) (*n*=357)	45.1%	33.9%	17.4%	3.6%
SN2. Your coworkers (i.e., other ECSE teachers and related service providers) (*n*=356)	24.7%	44.7%	27.8%	2.8%
SN3. Your students' parents (*n*=356)	27.5%	37.9%	31.2%	3.4%

SN4. How often, if ever, do you observe your coworkers (i.e., other ECSE teachers and related service providers) engaging in IEP data collection? (*n*=356)

- ○ Daily (4) – 16%
- ○ Weekly (3) – 39.6%
- ○ Monthly (2) – 23%
- ○ Less than monthly (1) – 21.4%

How often, if ever, do you observe someone in an ECSE leadership role (i.e., coordinator, program lead, professional development coach) in your district engage in these behaviors?

	Monthly (4)	Quarterly (3)	Yearly (2)	Never (1)
SN5. Communicate with you about IEP data collection (*n*=356)	15.4%	28.7%	25.3%	30.6%
SN6. Look at your IEP data (*n*=356)	10.1%	15.2%	19.4%	55.3%
SN7. Acknowledge you for your IEP data collection efforts (*n*=357)	7.6%	15.4%	23.5%	53.5%
SN8. Collect data highlighting student, staff, or program performance (*n*=356)	12.1%	19.4%	36.2%	32.3%

The following 4 questions relate to your beliefs about factors that impact your ability to engage in frequent IEP data collection.

How **difficult**, if at all, do you find these things?

	Not at all difficult (4)	Slightly difficult (3)	Mostly difficult (2)	Very difficult (1)
SED1. Writing clear and measurable IEP objectives across ALL developmental domains (*n*=338)	23.7%	58%	15.4%	2.9%
SED2. Modifying existing assessment tools to meet my IEP data collection needs (*n*=337)	13.7%	54.9%	23.7%	7.7%
SED3. Creating new assessment tools that meet my IEP data collection needs (*n*=337)	19.3%	43%	28.2%	9.5%
SED4. Designing a plan for the daily collection of IEP data (*n*=337)	21.1%	40.6%	25.8%	12.5%
SED5. Consistently carrying out my plan, such that IEP data are collected on a daily basis (*n*=336)	10.4%	36.3%	30.7%	22.6%

How **confident**, if at all, are you in your ability to do these things?

	Extremely confident (4)	Mostly confident (3)	Slightly confident (2)	Not at all confident (1)
SEC1. Write clear and measurable IEP objectives across ALL developmental domains? (*n*=336)	15.8%	60.7%	22.3%	1.2%
SEC2. Modify existing assessment tools to meet my IEP data collection needs? (*n*=336)	9.5%	48.7%	38.5%	3.3%
SEC3. Create new assessment tools that meet my IEP data collection needs? (*n*=336)	10.7%	41.4%	37.8%	10.1%
SEC4. Design a plan for the daily collection of IEP data? (*n*=335)	12.8%	46.7%	33.3%	7.2%
SEC5. Consistently carry out my plan, such that IEP data are collected on a daily basis? (*n*=335)	8.7%	35.8%	43.9%	11.6%

To what extent, if at all, do you <u>agree</u> with these statements?

	Strongly Agree (4)	Agree (3)	Disagree (2)	Strongly Disagree (1)
PBC1. The quality of my students' IEP objectives is entirely up to me. (*n*=335)	30.5%	51%	17.3%	1.2%
PBC2. The choice of tools I use to collect IEP data is entirely up to me. (*n*=335)	31.9%	49%	17%	2.1%
PBC3. Whether or not I engage in IEP data collection is entirely up to me. (*n*=333)	18%	31.5%	40.9%	9.6%

How often, if ever, do these factors **decrease your ability** to engage in <u>daily</u> IEP data collection?

	Never (4)	Sometimes (3)	Frequently (2)	Always (1)
B1. Availability of time to plan for IEP data collection (*n*=337)	3.3%	32.3%	43.6%	20.8%
B2. Availability of time to collect IEP data (*n*=337)	2.1%	30.6%	40.9%	26.4%
B3. Availability of time to analyze and interpret IEP data (*n*=335)	2.7%	28.3%	44.5%	24.5%
B4. Access to commercially-available paper-pencil tools that meet my IEP data collection needs (*n*=335)	40%	30.1%	20.9%	9%
B5. Access to commercially-available electronic tools that meet my IEP data collection needs (*n*=334)	29.3%	32.6%	24.6%	13.5%
B6. Classroom type (i.e., self-contained vs. inclusive) (*n*=334)	18.3%	38.3%	31.7%	11.7%
B7. Classroom curriculum (*n*=336)	22.9%	43.5%	26.2%	7.4%
B8. Nature of classroom routines/activities (i.e., balance of highly structured and adult-led one-on-one activities vs. minimally structured and child-led group activities) (*n*=336)	10.7%	47.6%	31.3%	10.4%
B9. Adult-to-child ratio in the classroom (*n*=335)	17%	45.4%	27.5%	10.1%

B10. Student's developmental levels in the classroom (*n*=336)	14.9%	49.7%	28.6%	6.8%
B11. Clarity of IEP objectives on your students' IEPs (*n*=334)	28.7%	57.5%	12.3%	1.5%
B12. Number of IEP objectives on your students' IEPs (*n*=337)	15.7%	51.1%	26.4%	6.8%
B13. Range of developmental levels across all students on your caseload (*n*=337)	11.9%	42.4%	36.2%	9.5%
B14. Number of students on your caseload (*n*=335)	5.1%	34.3%	37.9%	22.7%
B15. Your role in the planning and delivery of instruction and/or support within the classroom (*n*=331)	12.1%	40.2%	35.6%	12.1%
B16. Availability of professional development geared toward IEP data collection in ECSE (*n*=337)	14.5%	41.8%	30.9%	12.8%
B17. Other data collection requirements aligned with program, school, district, and/or state initiatives (*n*=337)	10.1%	47.2%	32%	10.7%

The following 3 questions relate to your <u>intentions to engage in future</u> IEP data collection.

I1. To what extent, if at all, do you plan to collect IEP data across ALL students on your caseload <u>over the next month</u>? (*n*=332)

- ○ To a great extent (4) – 47.6%
- ○ To a moderate extent (3) – 44.9%
- ○ To a slight extent (2) – 7.2%
- ○ Not at all (1) – 0.3%

I2. To what extent, if at all, is IEP data collection a high priority for you <u>over the next month</u>? (*n*=331)

- ○ To a great extent (4) – 44.1%
- ○ To a moderate extent (3) – 44.1%
- ○ To a slight extent (2) – 10.3%
- ○ Not at all (1) – 1.5%

Over the next month, to what extent, if at all, **do you intend to engage in** these forms of IEP data collection?

	To a great extent (4)	To a moderate extent (3)	To a slight extent (2)	Not at all (1)
I3. Reflecting on my memory of students' skills/behaviors while writing down informal notes (*n*=333)	40.8%	42.1%	15.9%	1.2%
I4. Talking with others (i.e., educational assistants, related service providers, and parents) about their observations of students' skills/behaviors while taking informal notes (*n*=333)	43.8%	46.3%	9.9%	0%
I5. Observing students while taking informal notes about their skills/behaviors (*n*=333)	49.9%	41.7%	8.1%	0.3%
I6. Observing students while formally documenting the frequency and/or duration of their skills/behaviors (*n*=333)	33%	38.5%	24.9%	3.6%
I7. Observing students while formally documenting the frequency and/or duration of their skills/behaviors AND the type and/or frequency of adult prompts provided (*n*=331)	26.6%	40.8%	25.7%	6.9%

Now, just a few questions about you!

What range best describes the **number of hours are you contracted to work each week**? (*n*=332)

- ○ Less than 10 (1) – 0.6%
- ○ 10-20 (2) – 3.3%
- ○ 21-30 (3) – 1.5%
- ○ 31-40 (4) – 94.6%

What range best describes the **number of students (birth through age 6) on your current caseload**? (*n*=332)

- ○ 1-5 (1) – 2.4%
- ○ 6-10 (2) – 9.6%
- ○ 10-15 (3) – 28.3%
- ○ 16-20 (4) – 43.4%
- ○ 21-25 (5) – 12.1%
- ○ More than 25 (6) – 4.2%

Approximately how often are you in a **classroom setting with at least one child on an IEP**? (*n*=332)

- ○ 5 days each week (1) – 62.4%
- ○ 4 days each week (2) – 22.9%
- ○ 3 days each week (3) – 5.7%
- ○ 2 days each week (4) – 5.7%
- ○ 1 day each week (5) – 2.1%
- ○ less than 1 day each week (6) – 1.2%

In the last month, how often, if at all, **did you engage in** these forms of IEP data collection?

	Frequently (4)	Sometimes (3)	Rarely (2)	Never (1)
Reflecting on my memory of students' skills/behaviors while writing down informal notes (*n*=331)	59.5%	32.9%	7%	0.6%
Talking with others (i.e., educational assistants, related service providers, and parents) about their observations of students' skills/behaviors while taking informal notes (*n*=331)	67.7%	30.2%	2.1%	0%
Observing students while taking informal notes about their skills/behaviors (*n*=331)	57.7%	36.6%	5.4%	0.3%
Observing students while formally documenting the frequency and/or duration of their skills/behaviors (*n*=331)	37.8%	42%	16.6%	3.6%
Observing students while formally documenting the frequency and/or duration of their skills/behaviors AND the type and/or frequency of adult prompts provided (*n*=329)	30.7%	41.9%	21.3%	6.1%

How many years have you been employed as a ***licensed ECSE teacher***? (*n*=331)

- ○ Less than 1 year (1) – 9.7%
- ○ 1-3 years (2) – 18.7%
- ○ 4-8 years (3) – 25.4%
- ○ 9-13 years (4) – 16%
- ○ 14-18 years (5) – 9.4%
- ○ 19-23 years (6) – 8.4%
- ○ More than 23 years (7) – 12.4%

Of those years, how many were spent working with **3 to 5-year-olds in classroom settings**? (*n*=331)

- ○ Less than 1 year (1) – 10.9%
- ○ 1-3 years (2) – 23%
- ○ 4-8 years (3) – 29.6%
- ○ 9-13 years (4) – 13.3%
- ○ 14-18 years (5) – 6.6%
- ○ 19-23 years (6) – 7.5%
- ○ More than 23 years (7) – 9.1%

Please refer to this map when answering the final question.

In which region are you currently employed as an ECSE teacher? (*n*=332)

- ○ 1: Kittson, Marshall, Norman, Pennington, Polk, Red Lake, or Roseau county (1) – 1.5%
- ○ 2: Beltrami, Clearwater, Hubbard, Lake of the Woods, or Mahnomen county (2) – 0.9%
- ○ 3: Aikin, Carton, Cook, Itasca, Koochiching, Lake, or St. Louis county (3) – 6.3%
- ○ 4: Becker, Clay, Douglas, Grant, Otter Tail, Pope, Stevens, Traverse, or Wilkin county (4) – 3.3%
- ○ 5: Cass, Crow Wing, Morrison, Todd, or Wadena county (5) – 3%
- ○ 6: Big Stone, Chippewa, Kandiyohi, Lac Qui Parle, Meeker, McLeod, Renville, Swift, or Yellow Medicine county (6) – 5.1%
- ○ 7: Benton, Chisago, Kanabec, Isanti, Mille Lacs, Pine, Sherburne, Sterns, or Wright county (7) – 10.8%
- ○ 8: Cottonwood, Jackson, Lincoln, Lyon, Murray, Nobles, Pipestone, Redwood, or Rock county (8) – 4.2%
- ○ 9: Blue Earth, Brown, Faribault, La Sueur, Martin, Nicollet, Sibley, Waseca, or Watonwan county (9) – 5.4%
- ○ 10: Dodge, Goodhue, Houston, Fillmore, Freeborn, Mower, Olmsted, Rice, Steele, Wabasha, or Winona county (10) – 7.5%
- ○ 11: Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, or Washington county (11) – 51.8%

If there is any additional information you want the researcher to know about IEP data collection in ECSE, please include it below.

____

If you'd like to be included in a drawing for one of 10 $25 Visa gift cards, enter your email address below.

**Please note that if you chose to enter the drawing, your email address will be immediately separated from the dataset, such that your survey responses remain anonymous. If you are chosen as a winner, you will receive an email from Brenna Rudolph (nolan225@umn.edu) with additional details.*

____